

# Binary quartic forms having bounded invariants, and the boundedness of the average rank of elliptic curves

Manjul Bhargava and Arul Shankar

June 10, 2010

## Abstract

We prove a theorem giving the asymptotic number of binary quartic forms having bounded invariants; this extends, to the quartic case, the classical results of Gauss and Davenport in the quadratic and cubic cases, respectively. Our techniques are quite general, and may be applied to counting integer orbits in other representations of algebraic groups. As an example, we illustrate the method on a particularly rich nonreductive representation having dimension 12.

We use these counting results to deduce a number of arithmetic consequences. First, we determine the mean number of 2-torsion elements in the class groups of *monogenic* maximal cubic orders. We similarly determine the mean number of 2-torsion elements in the class groups of maximal cubic orders having a monogenic subring of bounded index. Surprisingly, we find that these mean values are different! This demonstrates that, on average, the monogenicity of a ring has a direct altering effect on the behavior of the class group.

Finally, we utilize all the above results to prove that the average rank of elliptic curves, when ordered by their heights, is bounded. In particular, we prove that when elliptic curves are ordered by height, the mean size of the 2-Selmer group is 3. This implies that the average rank of elliptic curves is at most 1.5.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Average ranks of elliptic curves . . . . .	2
1.2	Counting binary forms having bounded invariants (particularly quartic forms) . . . . .	3
1.3	The mean size of the 2-class group of monogenic cubic fields . . . . .	7
<b>2</b>	<b>The number of integral binary quartic forms having bounded invariants</b>	<b>9</b>
2.1	Reduction theory . . . . .	10
2.2	Estimates on reducibility . . . . .	12
2.3	Averaging and cutting off the cusp . . . . .	13
2.4	Computing the volume . . . . .	15
2.5	Congruence conditions . . . . .	17
<b>3</b>	<b>The mean number of 2-torsion elements in the class groups of monogenic cubic fields</b>	<b>17</b>
3.1	The parametrization of monogenic cubic rings . . . . .	18
3.2	Counting monogenized cubic rings by height . . . . .	18
3.3	The parametrization of quartic rings having monogenic cubic resolvent rings . . . . .	20
3.4	The parametrization of 2-torsion elements in the class groups of monogenic cubic fields . . . . .	22
3.5	Computations of $p$ -adic densities . . . . .	23
3.6	Congruence conditions and uniformity estimates . . . . .	26
3.7	Proof of the main theorem (Theorem 1.9) . . . . .	27
3.8	Proofs of auxiliary lemmas . . . . .	29

<b>4</b>	<b>The mean number of 2-torsion elements in the class groups of submonogenized cubic fields</b>	<b>31</b>
4.1	The number of submonogenized cubic rings having bounded height and bounded index . . . .	31
4.1.1	The parametrization of submonogenized cubic rings of index $n$ . . . . .	31
4.1.2	Reduction theory . . . . .	32
4.1.3	Estimates on reducibility . . . . .	33
4.1.4	Averaging and cutting off the cusp . . . . .	33
4.1.5	Computing the volume . . . . .	34
4.2	The number of quartic rings having submonogenized cubic resolvent rings of bounded height and bounded index . . . . .	35
4.2.1	The parametrization of quartic rings having submonogenized cubic resolvent rings of index $n$ . . . . .	35
4.2.2	Reduction theory . . . . .	36
4.2.3	Estimates on reducibility . . . . .	38
4.2.4	Averaging and cutting off the cusp . . . . .	38
4.2.5	Computing the volume . . . . .	39
4.3	Congruence conditions . . . . .	40
4.4	Values of $p$ -adic densities . . . . .	40
4.5	Uniformity estimates . . . . .	41
4.6	Proof of Lemma 4.12 . . . . .	42
4.7	Proof of the main theorem (Theorem 1.10) . . . . .	42
<b>5</b>	<b>The mean size of the 2-Selmer group of elliptic curves</b>	<b>43</b>
5.1	Integral binary quartic forms and the 2-Selmer group of elliptic curves . . . . .	44
5.2	Computations of $p$ -adic densities in terms of local masses . . . . .	45
5.3	Uniformity estimates . . . . .	47
5.4	Proofs of the main theorems (Theorems 1.1, 1.3, and 5.1) . . . . .	47

# 1 Introduction

## 1.1 Average ranks of elliptic curves

Any elliptic curve  $E$  over  $\mathbb{Q}$  is isomorphic to a unique curve of the form  $E_{A,B} : y^2 = x^3 + Ax + B$ , where  $A, B \in \mathbb{Z}$  and for all primes  $p$ :  $p^6 \nmid B$  whenever  $p^4 \mid A$ . Let  $H(E_{A,B})$  denote the (naive) *height* of  $E_{A,B}$ , defined by  $H(E_{A,B}) := \max\{4|A^3|, 27B^2\}$ . Let  $\Delta(E_{A,B})$  and  $C(E_{A,B})$  denote the discriminant and conductor of  $E_{A,B}$ , respectively.

It is an old conjecture, originating in works of Goldfeld [29] and Katz-Sarnak [32], that a density of 50% of all elliptic curves over  $\mathbb{Q}$  have rank 0 and 50% have rank 1. These densities are expected to hold true regardless of whether one orders curves by height, discriminant, or conductor. In particular, one expects the average rank of all elliptic curves to be  $1/2$ . However, it has not previously been known that the average rank of all elliptic curves is even *finite* (i.e., bounded).

In [14], Brumer showed that the generalized Riemann hypothesis and the Birch–Swinnerton-Dyer conjectures together imply that the average rank of all elliptic curves, when ordered by their heights, is finite and is in fact bounded above by 2.3. Still assuming the generalized Riemann hypothesis and the Birch–Swinnerton-Dyer conjectures, this constant was subsequently improved to 2 by Heath-Brown [30] and to  $25/14 \sim 1.79$  by Young [42].

The purpose of this article is to prove unconditionally that the average rank of all elliptic curves, when ordered by their heights, is finite. In fact, we prove the same for the 2-Selmer rank. Recall that the 2-Selmer group  $S_2(E)$  of an elliptic curve fits into an exact sequence

$$0 \rightarrow E(\mathbb{Q})/2E(\mathbb{Q}) \rightarrow S_2(E) \rightarrow \text{III}_E[2] \rightarrow 0, \quad (1)$$

where  $\text{III}_E[2]$  denotes the 2-torsion subgroup of the Tate-Shafarevich group  $\text{III}_E$  of  $E$ . The 2-Selmer group is thus an elementary abelian 2-group, and its order is  $2^s$  for some integer  $s \geq 0$ ; the quantity  $s$  is called the *2-Selmer rank* of  $E$ .

Our main theorem on the 2-Selmer group is as follows:

**Theorem 1.1** *When all elliptic curves  $E/\mathbb{Q}$  are ordered by height, the average size of the 2-Selmer group  $S_2(E)$  is 3.*

We immediately conclude that:

**Corollary 1.2** *When all elliptic curves over  $\mathbb{Q}$  are ordered by height, their average 2-Selmer rank is at most 1.5; thus their average rank is also at most 1.5.*

Indeed, note that equation (1) implies that

$$r_2(S_2(E)) = r(E) + r_2(E[2]) + r_2(\text{III}_E[2]), \quad (2)$$

where we have used  $r(E)$  to denote the rank of  $E$  and  $r_2(G)$ , where  $G$  is an elementary abelian 2-group, to denote  $\dim_{\mathbb{F}_2}(G)$ . Theorem 1.1 bounds the left hand side of (2) by 1.5, and thus the same bound holds for the average size of each of the terms on the right hand side of (2). In particular, the average size of  $r_2(\text{III}_E[2])$  is also at most 1.5. Meanwhile, we will show in Section 5 that the mean size of  $r_2(E[2])$  is 0, i.e., 0% of elliptic curves possess rational 2-torsion.

We will in fact prove a stronger version of Theorem 1.1, namely:

**Theorem 1.3** *When elliptic curves  $E : y^2 = g(x)$ , in any family defined by finitely many congruence conditions on the coefficients of  $g$ , are ordered by height, the average size of the 2-Selmer group  $S_2(E)$  is 3.*

Thus the average size of the 2-Selmer group remains 3 even when one averages over any subset of elliptic curves defined by finitely many congruence conditions. We will actually prove Theorem 1.3 for an even larger class of families, including some that are defined by certain natural *infinite* sets of congruence conditions.

## 1.2 Counting binary forms having bounded invariants (particularly quartic forms)

We prove the above theorems by developing techniques to count integer orbits, having bounded invariants, in certain coregular representations over  $\mathbb{Z}$ . Recall that a *coregular representation* is a pair  $(G, V)$ , where  $G$  is an algebraic group and  $V$  is a representation of  $G$  (for our purposes, both defined over  $\mathbb{Z}$ ) such that the ring of polynomial invariants of  $G$  on  $V$  is free. Although our techniques are quite general, in this article we concentrate primarily on the case where  $G = \text{GL}_2$  and  $V$  is the space of **binary quartic forms**  $ax^4 + bx^3y + cx^2y^2 + dx^3y + ey^4$ .

The problem of counting integral binary forms having bounded invariants is a classical one. The case of binary quadratic forms was first treated in the influential work *Disquisitiones Arithmeticae* of Gauss in 1801. Gauss studied the action of  $\text{SL}_2(\mathbb{Z})$  on the space of integral binary quadratic forms  $f(x, y) = ax^2 + bxy + cy^2$  ( $a, b, c \in \mathbb{Z}$ )<sup>1</sup> via linear substitution of variable, in terms of the unique polynomial invariant for this action, namely the discriminant  $\Delta(f) = b^2 - 4ac$ . (The polynomial invariant  $\Delta(f)$  is “unique” in the sense that the ring of polynomial invariants is generated by one element, namely  $\Delta(f)$ .)

Gauss conjectured, and Mertens [36] and Siegel [39] proved, respectively, that:

**Theorem 1.4 (Gauss 1801/Mertens 1874/Siegel 1944)** *Let  $h_D$  denote the number of  $\text{SL}_2(\mathbb{Z})$ -equivalence classes of irreducible integral binary quadratic forms having discriminant  $D$ . Then:*

$$(a) \quad \sum_{-X < D < 0} h_D \sim \frac{\pi}{18} \cdot X^{3/2};$$

---

<sup>1</sup>Gauss actually considered only forms where  $b$  is even; however, from the modern point of view, it is natural to allow all three coefficients  $a, b, c$  to be arbitrary integers.

$$(b) \sum_{0 < D < X} h_D \log \varepsilon_D \sim \frac{\pi^2}{18} \cdot X^{3/2};$$

here  $\varepsilon_D = (t + u\sqrt{D})/2$ , where  $t, u$  are the smallest positive integral solutions of  $t^2 - du^2 = 4$ .

Note that  $h_D$  and  $\log \varepsilon_D$  have important algebraic number theoretic interpretations, namely,  $h(D)$  is the class number and  $\log \varepsilon_D$  is the regulator of the unique quadratic order of discriminant  $D$ . Thus Theorem 1.4(a) gives the average size of the class number of imaginary quadratic orders up to a given absolute discriminant, while (b) gives the average size of the class number times the regulator of real quadratic orders up to a given discriminant.

The next natural case to consider is that of integral binary cubic forms  $f(x, y) = ax^3 + bx^2y + cxy^2 + dy^3$  ( $a, b, c, d \in \mathbb{Z}$ ). The group  $\mathrm{GL}_2(\mathbb{Z})$  (or  $\mathrm{SL}_2(\mathbb{Z})$ ) again naturally acts on such forms, and there is again a unique polynomial invariant for this action, namely, the discriminant

$$\Delta(f) = b^2c^2 + 18abcd - 4ac^3 - 4b^3d - 27a^2d^2.$$

The question, as in the case of binary quadratic forms, is: how many classes  $h(D)$  of irreducible binary cubic forms are there with discriminant  $D$ , on average, as  $D$  varies?

This question was first answered by Davenport [23]:

**Theorem 1.5 (Davenport 1951)** *Let  $h(D)$  denote the number of  $\mathrm{GL}_2(\mathbb{Z})$ -equivalence classes of irreducible integral binary cubic forms having discriminant  $D$ . Then:*

$$(a) \sum_{-X < D < 0} h(D) \sim \frac{\pi^2}{24} \cdot X;$$

$$(b) \sum_{0 < D < X} h(D) \sim \frac{\pi^2}{72} \cdot X.$$

Davenport's theorem thus states that the number of equivalence classes of irreducible binary cubic forms per discriminant is a constant on average. This too has an important algebraic number theoretic interpretation. Since irreducible binary cubic forms are in bijection with orders in cubic fields (see Delone–Faddeev's work [26]), Theorem 1.5 states that there are a constant number of (isomorphism classes) of cubic orders per discriminant, on average. Davenport's Theorem was an essential ingredient in the classical work of Davenport and Heilbronn on the density of discriminants of cubic fields (see [24]).

The next natural case to consider is that of binary quartic forms. The group  $\mathrm{GL}_2(\mathbb{Z})$  again acts on the space of binary quartic forms  $f(x, y) = ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4$  ( $a, b, c, d, e \in \mathbb{Z}$ ) by linear substitution of variable. Note that in each of the cases of binary quadratic and binary cubic forms, the ring of invariants was generated by one element. Binary quartic forms historically have been more difficult to treat because the ring of invariants is now generated by two independent invariants, traditionally denoted  $I$  and  $J$ . For  $f(x, y)$  as above, we have the following explicit formulae for these invariants:

$$I(f) = 12ae - 3bd + c^2,$$

$$J(f) = 72ace + 9bcd - 27ad^2 - 27eb^2 - 2c^3.$$

Any other polynomial invariant for the action of  $\mathrm{GL}_2(\mathbb{Z})$  on binary quartic forms can be expressed as a polynomial in these invariants; for example, the discriminant  $\Delta(f)$  of a binary quartic form can be expressed in terms of  $I(f)$  and  $J(f)$  as follows:

$$\Delta(f) := \Delta(I, J) := (4I(f)^3 - J(f)^2)/27.$$

It follows from work of Borel and Harish-Chandra [13] that the number of equivalence classes of integral binary quartic forms, having any given fixed values of  $I$  and  $J$  (so long as both  $I$  and  $J$  are not

both equal to zero), is finite.<sup>2</sup> This raises the question as to how many classes  $h(I, J)$  of irreducible binary quartic forms with invariants  $I, J$  are there, on average, as the pair  $(I, J)$  varies?

To answer this question, we require just a bit of notation. Let us define the (naive) *height* of  $f(x, y)$  by  $H(f) := H(I, J) := \max\{|I|^3, J^2/4\}$  (the constant  $1/4$  on  $J^2$  is present for convenience, and is not of any real importance). Thus  $H(f)$  is a “degree 6” function on the coefficients of  $f$ , in the sense that  $H(rf) = r^6 H(f)$  for any constant  $r$ . We prove:

**Theorem 1.6** *Let  $h^{(i)}(I, J)$  denote the number of  $\mathrm{GL}_2(\mathbb{Z})$ -equivalence classes of irreducible binary quartic forms having  $4 - 2i$  real roots in  $\mathbb{P}^1$  and invariants equal to  $I$  and  $J$ . Then:*

$$\begin{aligned} \text{(a)} \quad \sum_{H(I, J) < X} h^{(0)}(I, J) &= \frac{4}{135} \zeta(2) X^{5/6} + O(X^{3/4+\epsilon}); \\ \text{(b)} \quad \sum_{H(I, J) < X} h^{(1)}(I, J) &= \frac{32}{135} \zeta(2) X^{5/6} + O(X^{3/4+\epsilon}); \\ \text{(c)} \quad \sum_{H(I, J) < X} h^{(2)}(I, J) &= \frac{8}{135} \zeta(2) X^{5/6} + O(X^{3/4+\epsilon}). \end{aligned}$$

In order to obtain the average size of  $h^{(i)}(I, J)$ , as  $(I, J)$  varies, we first wish to know which pairs  $(I, J)$  can actually occur as the invariants of an integral binary quartic form. In the quadratic and cubic cases, this is easy and well-known: a number occurs as the discriminant of a binary quadratic (resp. cubic) form if and only if it is congruent to 0 or 1 (mod 4).

In the binary quartic case, we prove that a similar scenario occurs, namely, an  $(I, J)$  is *eligible*—i.e., it occurs as the invariants of some integer binary quartic form—if and only if it satisfies any one of a certain specified finite set of congruence conditions modulo 27. More precisely, we prove:

**Theorem 1.7** *A pair  $(I, J) \in \mathbb{Z} \times \mathbb{Z}$  occurs as the invariants of an integral binary quartic form if and only if it satisfies one of the following congruence conditions:*

- (a)  $I \equiv 0 \pmod{3}$  and  $J \equiv 0 \pmod{27}$ ,
- (b)  $I \equiv 1 \pmod{9}$  and  $J \equiv \pm 2 \pmod{27}$ ,
- (c)  $I \equiv 4 \pmod{9}$  and  $J \equiv \pm 16 \pmod{27}$ ,
- (d)  $I \equiv 7 \pmod{9}$  and  $J \equiv \pm 7 \pmod{27}$ .

It follows that the number of eligible  $(I, J)$ , with  $H(I, J) < X$ , is a constant times  $X^{5/6}$ ; thus, by Theorem 1.6, the number of classes of binary quartic forms per eligible  $(I, J)$  is a constant on average. We have the following theorem:

**Theorem 1.8** *Let  $h^{(i)}(I, J)$  denote the number of  $\mathrm{GL}_2(\mathbb{Z})$ -equivalence classes of irreducible binary quartic forms having  $4 - 2i$  real roots and invariants equal to  $I$  and  $J$ . Let  $n_0 = 4$ ,  $n_1 = 2$ , and  $n_2 = 2$ . Then, for  $i = 0, 1, 2$ , we have:*

$$\lim_{X \rightarrow \infty} \frac{\sum_{H(I, J) < X} h^{(i)}(I, J)}{\sum_{\substack{(I, J) \text{ eligible} \\ (-1)^i \Delta(I, J) > 0 \\ H(I, J) < X}} 1} = \frac{2\zeta(2)}{n_i}.$$

---

<sup>2</sup>A deeper—and more ineffective—result of Birch and Merriman [11] states that the number of equivalence classes of binary quartic forms, having a fixed value of the single invariant  $\Delta(f) = 4I(f)^3 - J(f)^2$ , is finite. However, the latter fact will not be of importance to us here.

Thus, Theorem 1.8 says that the number of equivalence classes of binary quartic forms per eligible  $(I, J)$ , having a given number of real roots, is a constant on average. This constant is either  $\zeta(2)/2$  or  $\zeta(2)$ , depending on whether the given number of real roots is 4 or less than 4, respectively.

We in fact prove a strengthening of Theorem 1.8; namely, we obtain the asymptotic count of binary quartic forms, having bounded invariants, satisfying any specified finite set of congruence conditions. Such a modification will be crucial for several of the applications, which we now discuss.

First, we use the latter counting results involving binary quartic forms to understand the average size of 2-Selmer groups (as in Theorem 1.1). Recall that an element of the 2-Selmer group of an elliptic curve  $E/\mathbb{Q}$  may be thought of as a locally soluble 2-covering. A 2-covering of  $E/\mathbb{Q}$  is a genus one curve  $C$  together with maps  $\phi : C \rightarrow E$  and  $\theta : C \rightarrow E$ , where  $\phi$  is an isomorphism defined over  $\mathbb{C}$ , and  $\theta$  is a degree 4 map defined over  $\mathbb{Q}$ , such that the following diagram commutes:

$$\begin{array}{ccc} E & \xrightarrow{[2]} & E \\ \phi \uparrow & \nearrow \theta & \\ C & & \end{array}$$

Thus a 2-covering  $C = (C, \phi, \theta)$  may be viewed as a “twist over  $\mathbb{Q}$  of the multiplication-by-2 map on  $E$ ”. Two 2-coverings  $C$  and  $C'$  are said to be *isomorphic* if there exists an isomorphism  $\Phi : C \rightarrow C'$  defined over  $\mathbb{Q}$ , and a rational 2-torsion point  $P \in E$ , such that the following diagram commutes:

$$\begin{array}{ccc} E & \xrightarrow{+P} & E \\ \theta \uparrow & & \uparrow \theta' \\ C & \xrightarrow{\Phi} & C' \end{array}$$

A *soluble 2-covering*  $C$  is one that possesses a rational point, while a *locally soluble 2-covering*  $C$  is one that possesses an  $\mathbb{R}$ -point and a  $\mathbb{Q}_p$ -point for all primes  $p$ . Then we have, as groups, the isomorphisms:

$$\begin{aligned} \{\text{soluble 2-coverings}\} / \sim &\cong E(\mathbb{Q})/2E(\mathbb{Q}); \\ \{\text{locally soluble 2-coverings}\} / \sim &\cong S_2(E). \end{aligned}$$

How does counting elements of  $S_2(E)$  lead to counting binary quartic forms? There is a result of Birch and Swinnerton-Dyer (see [12, Lemma 2]) that states that any locally soluble 2-covering  $C$  possesses a degree 2 divisor defined over  $\mathbb{Q}$ , thus yielding a double cover  $C \rightarrow \mathbb{P}^1$  ramified at 4 points. We thus obtain a binary quartic form over  $\mathbb{Q}$ , well-defined up to  $\text{GL}_2(\mathbb{Q})$ -equivalence!

To carry out the proof of Theorems 1.1 and 1.3, we do the following:

- Given  $A, B \in \mathbb{Z}$ , choose an *integral* binary quartic form  $f$  for each element of  $S_2(E_{A,B})$ , such that
  - $y^2 = f(x)$  gives the desired 2-covering;
  - the invariants  $(I(f), J(f))$  of  $f$  agree with the invariants  $(A, B)$  of the elliptic curve (at least away from 2 and 3);
- Count these integral binary quartic forms via Theorem 1.6. The relevant binary quartic forms are defined by infinitely many congruence conditions, so a sieve has to be performed.
- A uniformity estimate must be proven to perform this sieve, and that is by far the most technical part of this work. It involves counting integral points, having bounded invariants, in a certain nonreductive coregular space of dimension 12.

To carry out the sieve/uniformity step in the arguments above, we require in particular some auxiliary results on the 2-class groups of *monogenic* cubic fields, which we believe are of interest in their own right. We describe these results next.

### 1.3 The mean size of the 2-class group of monogenic cubic fields

We define a *cubic ring* to be a commutative ring with unit that is free of rank 3 as a  $\mathbb{Z}$ -module. For example, an order in a cubic field is our prototypical cubic ring. We say that a cubic ring is *monogenic* if it is generated by a single element as a  $\mathbb{Z}$ -algebra. A *monogenized cubic ring* is a pair  $(C, x)$ , where  $C$  is a monogenic cubic ring and  $x \in C$  is an element that generates  $C$  over  $\mathbb{Z}$ . In that case,  $C$  is isomorphic to the ring  $\mathbb{Z}[X]/(X^3 + rX^2 + sX + t)$ , where  $f(X) = X^3 + rX^2 + sX + t$  is the characteristic polynomial of the linear map  $\times x : C \rightarrow C$  given by multiplication by  $x$ .

If  $(C, x)$  is a monogenized cubic ring, then so is  $(C, x - n)$  for any  $n \in \mathbb{Z}$ . We thus define an equivalence relation  $\sim$  on monogenized cubic rings by  $(C, x) \sim (C, y)$  if and only if  $x - y \in \mathbb{Z}$ . This leads to a corresponding translation action of  $\mathbb{Z}$  on monic cubic polynomials, namely  $n \in \mathbb{Z}$  sends  $f(X)$  to  $g(X) = f(X + n)$ , and  $\mathbb{Z}[X]/f(X) \sim \mathbb{Z}[X]/g(X)$ .

We show that this translation action of  $\mathbb{Z}$  on monic cubic polynomials  $f(X) = X^3 + rX^2 + sX + t$  again has two polynomial invariants, having degrees 2 and 3 respectively, and which we again call  $I(f)$  and  $J(f)$ . They are defined by

$$I(f) := r^2 - 3s,$$

$$J(f) := -2r^3 + 9rs - 27t.$$

One easily checks that these quantities do not change if  $f(X)$  is replaced by  $f(X + n)$  for any  $n \in \mathbb{Z}$ . We define the (naive) *height*  $H(C, x)$  of the monogenized cubic ring  $(C, x)$  by  $H(C, x) := H(I(f), J(f)) = \max\{|I(f)|^3, J(f)^2/4\}$ , where  $f$  as before denotes the characteristic polynomial of  $\times x : C \rightarrow C$ .

We will be concerned primarily with *monogenic cubic fields*, i.e., cubic fields whose rings of integers are monogenic. We say that  $(K, x)$  is a *monogenized cubic field* if  $(\mathcal{O}_K, x)$  is a monogenized cubic ring; here  $\mathcal{O}_K$  denotes the ring of integers in  $K$ . In that case, we define the *height* of the monogenized cubic field  $(K, x)$  by  $H(K) := H(\mathcal{O}_K, x)$ .

We now order all monogenized cubic fields  $K = (K, x)$  by their heights, and ask: what is the average size of the 2-torsion subgroups of the class groups of monogenized cubic fields? (It is known that a cubic field can be monogenized in at most 12 ways, up to equivalence; see [25] and [27].) We prove the following theorem:

**Theorem 1.9** *For a cubic field  $K$ , let  $\text{Cl}_2(K)$  (resp.  $\text{Cl}_2^+(K)$ ) denote the 2-torsion subgroup of the class group of  $K$  (resp. the narrow class group of  $K$ ). Let  $K$  run through all isomorphism classes of monogenized cubic fields ordered by height. Then:*

- (a) *The average size of  $\text{Cl}_2(K)$  over totally real monogenized cubic fields  $K$  is  $3/2$ ;*
- (b) *The average size of  $\text{Cl}_2(K)$  over complex monogenized cubic fields  $K$  is  $2$ ;*
- (c) *The average size of  $\text{Cl}_2^+(K)$  over totally real monogenized cubic fields  $K$  is  $5/2$ .*

*Moreover, these averages do not change even if  $K$  runs only over those monogenized cubic fields satisfying any specified set of local conditions at a finite set of primes.*

It is interesting to note that the averages in Theorem 1.9 are strictly *larger* than the corresponding averages for all cubic fields, when ordered by discriminant; indeed, the averages in the latter case are  $5/4$ ,  $3/2$ , and  $2$  respectively (see [5, Theorem 5]). This indicates that monogenic cubic fields tend to have more 2-torsion in their class groups, on average, than general cubic fields! However, a priori, it is also possible that this discrepancy arises because we are ordering these two sets of cubic fields differently (the first by height, and the second by discriminant).

We thus prove a result on the average number of 2-torsion elements in the class groups of general cubic fields, when they are ordered by *height* rather than discriminant. To make this precise, let us say that  $(C, x)$  is a *submonogenized cubic ring of index  $n$* , or simply an  *$n$ -monogenized cubic ring*, if  $C$  is a cubic ring and  $x \in C$  is an element such that  $x$  generates a subring of index  $n$  in  $C$ . Thus a 1-monogenized cubic ring is equivalent to a monogenized cubic ring. In general, an  $n$ -monogenized cubic ring is a cubic ring having an

index  $n$  subring that is monogenic, together with a choice of generator for this subring. Note that any cubic ring can be  $n$ -monogenized for some integer  $n \geq 1$ .

We define the *height*  $H(C, x)$  of an  $n$ -monogenized cubic ring  $(C, x)$  by  $H(C, x) := n^2 H(I(f), J(f)) = n^2 \max\{|I(f)|^3, J(f)^2/4\}$ , where  $f$  as before denotes the characteristic polynomial of  $\times x : C \rightarrow C$ . (The factor of  $n^2$  appears so that the height  $H(C, x)$  of  $(C, x)$  remains comparable with the discriminant  $\Delta(C)$  of  $C$ .)

Finally, an  *$n$ -monogenized cubic field* is a pair  $(K, x)$  for which  $(\mathcal{O}_K, x)$  is an  $n$ -monogenized cubic ring. We define the *height*  $H(K, x)$  of an  $n$ -monogenized cubic field  $(K, x)$  by  $H(K, x) := H(\mathcal{O}_K, x)$ .

We may now ask: what is the average size of the 2-torsion subgroup of the class group of  $n$ -monogenized cubic fields, over all isomorphism classes of  $n$ -monogenized fields having height less than  $X$  and  $n < X^\delta$ ? We prove the following theorem:

**Theorem 1.10** *For a cubic field  $K$ , let  $\text{Cl}_2(K)$  (resp.  $\text{Cl}_2^+(K)$ ) denote the 2-torsion subgroup of the class group of  $K$  (resp. the narrow class group of  $K$ ). Let  $\delta \leq 1/4$  be any positive constant, and let  $S(X, \delta)$  denote the set of all isomorphism classes of  $n$ -monogenized cubic fields  $K$  such that  $H(K) < X$  and  $n < X^\delta$ . Then:*

- (a) *As  $X \rightarrow \infty$ , the average size of  $\text{Cl}_2(K)$  over totally real cubic fields  $K$  in  $S(X, \delta)$  approaches  $5/4$ ;*
- (b) *As  $X \rightarrow \infty$ , the average size of  $\text{Cl}_2(K)$  over complex cubic fields  $K$  in  $S(X, \delta)$  approaches  $3/2$ ;*
- (c) *As  $X \rightarrow \infty$ , the average size of  $\text{Cl}_2^+(K)$  over totally real cubic fields  $K$  in  $S(X, \delta)$  approaches 2.*

*Moreover, these averages do not change even if  $K$  runs only over those  $n$ -monogenized cubic fields satisfying any specified set of local conditions at a finite set of primes.*

Theorem 1.10 makes it clear that, in a precise sense, the average size of the 2-torsion subgroups of class groups of monogenic cubic fields is *higher* than the corresponding average for general cubic fields! The theorem implies that the monogenicity of the ring of integers of a cubic field has a direct influence on the behaviour of the class group of the field. In particular, the class groups of monogenic cubic fields do not seem to be “random groups” in the sense of Cohen–Lenstra–Martinet [17] and Malle [35].

It would be interesting to have heuristics that explain this phenomenon. In a forthcoming paper, in joint work with Jonathan Hanke, we describe the average size  $s(n)$  of  $\text{Cl}_2(K)$  over  $n$ -monogenic cubic fields  $K$ , where  $n$  is *fixed*; the base case  $n = 1$  is Theorem 1.9. The function  $s(n)$  is found to vary in a very interesting way with  $n$ , and the average of  $s(n)$  over all  $n$  is then shown to yield the constants in Theorem 1.10; see [8].

We prove Theorem 1.9 by reducing it to an appropriate count of binary quartic forms satisfying a certain natural infinite set of congruence conditions. The key observation in this context is that  $\text{GL}_2(\mathbb{Z})$ -equivalence classes of integral binary quartic forms parametrize quartic rings having monogenic cubic resolvent rings (in the sense of [4]); this is a theorem of Wood [41, Theorem 4.1.1]. When we restrict to the binary quartic forms yielding *maximal* cubic resolvent rings, then by class field theory, such binary quartic forms may then be seen to correspond to the 2-torsion elements in the dual of the class group of monogenic cubic fields. By counting these binary quartic forms, we obtain Theorem 1.9.

However, since the set of binary quartic forms that must be counted is defined by infinitely many congruence conditions, a sieve must be performed to obtain the correct count. In fact, it is essentially the same sieve required to average the number of elements in the 2-Selmer group of elliptic curves (as in Theorems 1.1 and 1.3)!

While the sieve that we use is very simple, the key ingredient that is required to perform the sieve is a certain uniformity estimate on the error, as more and more congruence conditions are introduced, that is quite challenging and technical to prove. This uniformity estimate is in some sense the analogue of the uniformity estimate in Davenport–Heilbronn’s work [24, Proposition 1] (indeed, this is perhaps the most technical part of their paper) and of those in [5, Proposition 23] and [6, Proposition 18].

The case at hand seems considerably more difficult. We prove the desired estimate by enlarging the space of binary quartic forms, equipped with the action of  $\text{GL}_2(\mathbb{Z})$ , to the space of pairs of ternary quadratic forms, equipped with the action of a certain non-reductive group; this corresponds to embedding



the parameter space of all monogenized cubic rings into that of all submonogenized cubic rings. The uniformity estimate is then proven by counting orbits on this enlarged space, in conjunction with a class field theory argument that then applies in this context. Although this takes some effort, as a by-product and consequence of this work we end up proving Theorem 1.10!

This paper is organized as follows. In Section 2, we study the distribution of  $\mathrm{GL}_2(\mathbb{Z})$ -equivalence classes of binary quartic forms with respect to their fundamental invariants  $I$  and  $J$ ; in particular, we prove Theorems 1.6 and 1.8. In Section 3, we describe the correspondence between binary quartic forms and 2-class groups of monogenic cubic fields, by generalizing in a natural way the correspondence of Wood [41]. We use this to prove Theorem 1.9 (assuming the aforementioned uniformity estimates).

In Section 4, we prove the uniformity estimates that are necessary to count binary quartic forms satisfying our desired infinite sets of congruence conditions. In particular, we complete the proof of Theorem 1.9, and also obtain Theorem 1.10.

In Section 5, we conclude by describing the precise connection between binary quartic forms and elements in the 2-Selmer groups of elliptic curves. This connection allows us, through the use of certain mass formulae for elliptic curves over  $\mathbb{Q}_p$ , to compute the average size of the 2-Selmer groups of elliptic curves (or of appropriate families of elliptic curves) via a count of binary quartic forms satisfying a certain corresponding infinite set of congruence conditions. We then apply the uniformity results of Section 4 to count these binary quartic forms, thus completing the proofs of Theorems 1.1 and 1.3.

## 2 The number of integral binary quartic forms having bounded invariants

In this section, we derive precise asymptotics for the number of  $\mathrm{GL}_2(\mathbb{Z})$ -equivalence classes of irreducible integral binary quartic forms having bounded invariants. We also describe how these asymptotics change when we restrict to counting only those binary quartic forms satisfying any specified finite set of congruence conditions. In particular, we prove Theorems 1.6–1.8.

To this end, let  $V_{\mathbb{R}}$  denote the vector space of binary quartic forms over the real numbers  $\mathbb{R}$ . We express an element  $f \in V_{\mathbb{R}}$  in the form  $f(x, y) = ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4$ , where  $a, b, c, d$ , and  $e$  are real numbers. Such an  $f \in V_{\mathbb{R}}$  is said to be *integral* if  $a, b, c, d, e \in \mathbb{Z}$ .

The group  $\mathrm{GL}_2(\mathbb{R})$  naturally acts on  $V_{\mathbb{R}}$ ; namely, an element  $\gamma \in \mathrm{GL}_2(\mathbb{R})$  acts on  $f(x, y)$  by linear substitution of variable:

$$\gamma \cdot f(x, y) = f((x, y) \cdot \gamma). \quad (3)$$

This action of  $\mathrm{GL}_2(\mathbb{R})$  on  $V_{\mathbb{R}}$  is a left action, i.e.,  $(\gamma_1 \gamma_2) \cdot f = \gamma_1 \cdot (\gamma_2 \cdot f)$ .

We consider the action of  $\mathrm{SL}_2^{\pm}(\mathbb{R})$  on  $V_{\mathbb{R}}$ , where  $\mathrm{SL}_2^{\pm}(\mathbb{R}) \subset \mathrm{GL}_2(\mathbb{R})$  is the group of elements in  $\mathrm{GL}_2(\mathbb{R})$  having determinant equal to  $\pm 1$ . The ring of invariants for this action is generated by two independent generators of degrees 2 and 3 which are traditionally denoted by  $I$  and  $J$ , respectively. If  $f(x, y) = ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4$ , then

$$\begin{aligned} I(f) &= 12ae - 3bd + c^2, \\ J(f) &= 72ace + 9bcd - 27ad^2 - 27eb^2 - 2c^3. \end{aligned} \quad (4)$$

The quantities  $I(f)$  and  $J(f)$  are also *relative invariants* for the action of  $\mathrm{GL}_2(\mathbb{R})$  on  $V$ : we have  $I(\gamma \cdot f) = (\det \gamma)^4 I(f)$  and  $J(\gamma \cdot f) = (\det \gamma)^6 J(f)$ . The discriminant  $\Delta(f)$  of a binary quartic form  $f$ , being a relative invariant of degree 6, can thus be expressed in terms of  $I$  and  $J$ , namely,  $\Delta(f) = (4I^3(f) - J^2(f))/27$ . We define the *height*  $H(f)$  of a binary quartic form  $f$  by:

$$H(f) := H(I, J) = \max\{|I|^3, J^2/4\}. \quad (5)$$

The action of  $\mathrm{GL}_2(\mathbb{Z})$  on  $V_{\mathbb{R}}$  evidently preserves the lattice  $V_{\mathbb{Z}}$  consisting of the integral elements of  $V_{\mathbb{R}}$ , and so we may ask: how many  $\mathrm{GL}_2(\mathbb{Z})$ -classes of forms are there having height at most  $X$ ? More precisely, we may ask: how many  $\mathrm{GL}_2(\mathbb{Z})$ -classes of forms are there with height at most  $X$  and a given

number of real roots? To this end, for  $i = 0, 1$ , and  $2$ , let  $V_{\mathbb{Z}}^{(i)}$  denote the set of elements in  $V_{\mathbb{Z}}$  having  $4 - 2i$  real roots in  $\mathbb{P}^1(\mathbb{R})$ . Then the main theorem of this section is the following restatement of Theorem 1.6:

**Theorem 2.1** *For any  $\mathrm{GL}_2(\mathbb{Z})$ -invariant set  $S \subset V_{\mathbb{Z}}$ , let  $N(S; X)$  denote the number of  $\mathrm{GL}_2(\mathbb{Z})$ -equivalence classes of irreducible elements  $f \in S$  satisfying  $H(f) < X$ . Then*

$$(a) \quad N(V_{\mathbb{Z}}^{(0)}; X) = \frac{4}{135}\zeta(2)X^{5/6} + O(X^{3/4+\epsilon});$$

$$(b) \quad N(V_{\mathbb{Z}}^{(1)}; X) = \frac{32}{135}\zeta(2)X^{5/6} + O(X^{3/4+\epsilon});$$

$$(c) \quad N(V_{\mathbb{Z}}^{(2)}; X) = \frac{8}{135}\zeta(2)X^{5/6} + O(X^{3/4+\epsilon}).$$

Our strategy to prove Theorem 2.1 is as follows. In §2.1, we develop the necessary reduction theory needed to establish convenient fundamental domains for the action of  $\mathrm{GL}_2(\mathbb{Z})$  on  $V_{\mathbb{R}}$ . The primary difficulty in counting points in these fundamental domains is that they are not compact, but instead have a rather complicated cuspidal region going off to infinity. To deal with and effectively handle this cusp, in §2.2 we investigate the distribution of reducible and irreducible points inside these fundamental domains. Specifically, we prove that the cusp contains only reducible points, while the remainder of the domain outside the cuspidal region contains primarily irreducible points. In §2.3, we develop a refinement of an averaging method introduced in [5], [6] to count points in these fundamental regions in terms of the volumes of these domains. The volumes of the fundamental regions are then computed in §2.4, completing the proof of Theorem 2.1.

Finally, in §2.5, we prove a stronger version of Theorem 2.1 where we restrict to counting those binary quartic forms whose coefficients satisfy finitely many congruence conditions. We will require this stronger result in an essential way when we prove Theorems 1.1, 1.3, and 1.9 in Sections 5, 5, and 3, respectively.

## 2.1 Reduction theory

For  $i = 0, 1$ , and  $2$ , let  $V_{\mathbb{R}}^{(i)}$  denote the set of points in  $V_{\mathbb{R}}$  having exactly  $4 - 2i$  real roots in  $\mathbb{P}^1$ . Then it is easy to see that  $V_{\mathbb{R}}^{(2)}$  is the set of *definite* forms in  $V_{\mathbb{R}}$ , i.e., forms  $f(x, y)$  that take only positive or only negative values when evaluated at nonzero vectors  $(x_0, y_0) \in \mathbb{R}^2$ . Let  $V_{\mathbb{R}}^{(2+)}$  (resp.  $V_{\mathbb{R}}^{(2-)}$ ) denote the subset of  $V_{\mathbb{R}}^{(2)}$  consisting of the *positive definite forms* (resp. *negative definite forms*). Note that for  $i = 0, 1$ , and  $2$  we have  $V_{\mathbb{Z}}^{(i)} = V_{\mathbb{R}}^{(i)} \cap V_{\mathbb{Z}}$ . We analogously define  $V_{\mathbb{Z}}^{(i)} = V_{\mathbb{R}}^{(i)} \cap V_{\mathbb{Z}}$  for  $i = 2+$  and  $2-$ .

We have the following facts (see [20, Remark 2]):

1. The set of binary quartic forms in  $V_{\mathbb{R}}$  having fixed invariants  $I$  and  $J$  consists of just one  $\mathrm{GL}_2(\mathbb{R})$ -orbit if  $4I^3 - J^2 < 0$ ; this orbit lies in  $V_{\mathbb{R}}^{(1)}$ .
2. The set of binary quartic forms in  $V_{\mathbb{R}}$  having fixed invariants  $I$  and  $J$  consists of three  $\mathrm{GL}_2(\mathbb{R})$ -orbits if  $4I^3 - J^2 > 0$ ; in that case, there is one such orbit from each of  $V_{\mathbb{R}}^{(0)}$ ,  $V_{\mathbb{R}}^{(2+)}$ , and  $V_{\mathbb{R}}^{(2-)}$ .

We construct fundamental sets  $L_V^{(i)}$  for the action of  $\mathrm{GL}_2(\mathbb{R})$  on  $V_{\mathbb{R}}^{(i)}$  for each  $i \in \{0, 1, 2+, 2-\}$ . Let us first consider the case where  $i$  is  $0, 2+$ , or  $2-$ . In that case, if a form  $f \in V_{\mathbb{R}}^{(i)}$  has invariants  $I$  and  $J$ , then  $4I^3 - J^2 > 0$ ; furthermore,  $f$  is then the unique form in  $V_{\mathbb{R}}^{(i)}$  up to  $\mathrm{GL}_2(\mathbb{R})$ -equivalence having invariants  $I$  and  $J$ . Similarly, if  $f \in V_{\mathbb{R}}^{(1)}$  has invariants  $I$  and  $J$ , then  $4I^3 - J^2 < 0$  and  $f$  is then the unique form in  $V_{\mathbb{R}}$ , up to  $\mathrm{GL}_2(\mathbb{R})$ -equivalence, having invariants  $I$  and  $J$ .

Since  $I(g \cdot f) = (\det g)^4 I(f)$  and  $J(g \cdot f) = (\det g)^6 J(f)$ , it is clear that two forms  $f_1, f_2 \in V_{\mathbb{R}}^{(i)}$  are  $\mathrm{GL}_2(\mathbb{R})$ -equivalent if and only if there exists a positive constant  $\lambda \in \mathbb{R}$  with  $I(f_1) = \lambda^2 I(f_2)$  and  $J(f_1) = \lambda^3 J(f_2)$ . In particular, given a pair  $(I, J) \neq (0, 0)$ , there always exists a positive constant  $\lambda$  such that  $H(\lambda^2 I, \lambda^3 J) = 1$ . It follows that, for  $i = 0, 2+$ , or  $2-$  (resp. for  $i = 1$ ), a fundamental set  $L_V^{(i)}$  for the action of  $\mathrm{GL}_2(\mathbb{R})$  on  $V_{\mathbb{R}}^{(i)}$  can be constructed by choosing one form  $f \in V_{\mathbb{R}}^{(i)}$  for each  $(I, J)$  such that

$$\begin{aligned}
L_V^{(0)} &= \left\{ x^3y - \frac{1}{3}xy^3 - \frac{t}{27}y^4 : -2 \leq t \leq 2 \right\} \\
L_V^{(1)} &= \left\{ x^3y - \frac{s}{3}xy^3 + \frac{\pm 2}{27}y^4 : -1 \leq s \leq 1 \right\} \cup \left\{ x^3y + \frac{1}{3}xy^3 - \frac{t}{27}y^4 : -2 \leq t \leq 2 \right\} \\
L_V^{(2+)} &= \left\{ \frac{1}{16}x^4 - \frac{\sqrt{2-t}}{3\sqrt{3}}x^3y + \frac{1}{2}x^2y^2 + y^4 : -2 \leq t \leq 2 \right\} \\
L_V^{(2-)} &= \{ f : -f \in L_V^{(2+)} \}
\end{aligned}$$

Table 1: Explicit constructions of fundamental sets  $L_V^{(i)}$  for  $\mathrm{GL}_2(\mathbb{R}) \backslash V_{\mathbb{R}}^{(i)}$

$H(I, J) = 1$  and  $4I^3 - J^2 > 0$  (resp.  $4I^3 - J^2 < 0$ ). Table 1 provides explicit constructions of such fundamental sets  $L_V^{(i)}$ . The key fact that we use about these chosen fundamental sets  $L_V^{(i)}$  is that the coefficients of all binary quartic forms in these  $L_V^{(i)}$  are *uniformly bounded*; i.e., the  $L_V^{(i)}$  all lie in a compact subset of  $V_{\mathbb{R}}$ . Note that for any fixed  $h \in \mathrm{GL}_2(\mathbb{R})$  (or, indeed, for any  $h$  lying in a fixed compact subset  $G_0 \subset \mathrm{GL}_2(\mathbb{R})$ ), the set  $hL_V^{(i)}$  is also a fundamental set for the action of  $\mathrm{GL}_2(\mathbb{R})$  on  $V_{\mathbb{R}}^{(i)}$ , and again all coefficients are uniformly bounded.

We will have need for the following lemma, whose proof is postponed to §3.8:

**Lemma 2.2** *Let  $f$  be an element in  $V_{\mathbb{R}}^{(i)}$  having nonzero discriminant. Then the order of the stabilizer of  $f$  in  $\mathrm{GL}_2(\mathbb{R})$  is 8 if  $i = 0$  or  $2$ , and 4 if  $i = 1$ .*

Let  $\mathcal{F}$  denote Gauss's usual fundamental domain for  $\mathrm{GL}_2(\mathbb{Z}) \backslash \mathrm{GL}_2(\mathbb{R})$  in  $\mathrm{GL}_2(\mathbb{R})$ . Then  $\mathcal{F}$  may be expressed in the form  $\mathcal{F} = \{ n\alpha k\lambda : n(u) \in N'(t), \alpha(t) \in A', k \in K, \lambda \in \Lambda \}$ , where

$$N'(\alpha) = \left\{ \begin{pmatrix} 1 & \\ u & 1 \end{pmatrix} : u \in \nu(\alpha) \right\}, \quad A' = \left\{ \begin{pmatrix} t^{-1} & \\ & t \end{pmatrix} : t \geq \sqrt{3}/2 \right\}, \quad \Lambda = \left\{ \begin{pmatrix} \lambda & \\ & \lambda \end{pmatrix} : \lambda > 0 \right\}, \quad (6)$$

and  $K$  is as usual the (compact) real orthogonal group  $\mathrm{SO}_2(\mathbb{R})$ ; here  $\nu(\alpha)$  is a union of one or two subintervals of  $[-\frac{1}{2}, \frac{1}{2}]$  depending only on the value of  $\alpha \in A'$ .

For  $i = 0, 1, 2+$ , and  $2-$ , let  $2 \cdot n_i$  denote the cardinality of the stabilizer in  $\mathrm{GL}_2(\mathbb{R})$  of an irreducible element  $v \in V_{\mathbb{R}}^{(i)}$ . Then, by Lemma 2.2, we have  $n_0 = 4$ ,  $n_1 = 2$ ,  $n_{2+} = 4$ , and  $n_{2-} = 4$ . Thus for any  $h \in \mathrm{GL}_2(\mathbb{R})$ , we see that  $\mathcal{F}hL_V^{(i)}$  is essentially the union of  $n_i$  fundamental domains for the action of  $\mathrm{GL}_2(\mathbb{Z})$  on  $V_{\mathbb{R}}^{(i)}$ ; here, we regard  $\mathcal{F}hL_V^{(i)}$  as a multiset, where the multiplicity of a point  $x$  in  $\mathcal{F}hL_V^{(i)}$  is given by the cardinality of the set  $\{g \in \mathcal{F} : x \in ghL_V^{(i)}\}$ . Thus a  $\mathrm{GL}_2(\mathbb{Z})$ -equivalence class  $x$  in  $V_{\mathbb{R}}^{(i)}$  is represented in this multiset  $m(x)$  times, where  $m(x) = \#\mathrm{Stab}_{\mathrm{GL}_2(\mathbb{R})}(x) / \#\mathrm{Stab}_{\mathrm{GL}_2(\mathbb{Z})}(x)$ . Since the stabilizer in  $\mathrm{GL}_2(\mathbb{Z})$  of an element  $x \in V_{\mathbb{Z}}$  always contains the identity and its negative,  $m(x)$  is always a number between 1 and  $n_i$ .

We conclude that, for any  $h \in \mathrm{GL}_2(\mathbb{R})$ , the product  $n_i \cdot N(V_{\mathbb{Z}}^{(i)}; X)$  is exactly equal to the number of irreducible integer points in  $\mathcal{F}hL_V^{(i)}$  having absolute discriminant less than  $X$ , with the slight caveat that the (relatively rare—see Lemma 2.4) points with  $\mathrm{GL}_2(\mathbb{Z})$ -stabilizers of cardinality  $2r$  ( $r > 1$ ) are counted with weight  $1/r$ .

As mentioned earlier, the main obstacle to counting integer points of bounded discriminant in a single domain  $\mathcal{F}hL_V^{(i)}$  is that the relevant region is not compact, but rather has a cusp going off to infinity. We simplify the counting in this cuspidal region by “thickening” the cusp; more precisely, we compute the number of integer points of bounded discriminant in the region  $\mathcal{F}hL_V^{(i)}$  by averaging over lots of such fundamental regions, i.e., by averaging over the domains  $\mathcal{F}hL_V^{(i)}$  where  $h$  ranges over a certain compact subset  $G_0 \in \mathrm{GL}_2(\mathbb{R})$ . This refinement of the method of [6] is described in more detail in §2.3.

However, we first turn in §2.2 to estimating the number of reducible points in the main bodies (i.e., away from the cusps) of our fundamental regions.

## 2.2 Estimates on reducibility

We consider the integral elements in the region  $\mathcal{R}_X(hL_V^{(i)}) := \{w \in \mathcal{F}hL_V^{(i)} : |H(w)| < X\}$  that are reducible over  $\mathbb{Q}$ , where  $h$  is any element in a fixed compact subset  $G_0$  of  $\mathrm{GL}_2(\mathbb{R})$ . Note that if a binary quartic form  $ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4$  satisfies  $a = 0$ , then it is automatically reducible over  $\mathbb{Q}$ , since  $y$  is a factor. The following lemma shows that for integral binary quartic forms in  $\mathcal{R}_X(hL_V^{(i)})$ , reducibility with  $a \neq 0$  does not occur very often:

**Lemma 2.3** *Let  $h \in G_0$  be any element, where  $G_0$  is any fixed compact subset of  $\mathrm{GL}_2(\mathbb{R})$ . Then the number of integral binary quartic forms  $ax^4 + bx^3y + cy^2 + dxy^3 + ey^4 \in \mathcal{R}_X(hL_V^{(i)})$  that are reducible over  $\mathbb{Q}$  with  $a \neq 0$  is  $O(X^{2/3+\epsilon})$ , where the implied constant depends only on  $G_0$ .*

**Proof:** Let  $f(x, y) = ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4$  be any element in  $\mathcal{R}_X(hL_V^{(i)})$ . Then since  $\mathcal{R}_X(hL_V^{(i)}) \subset N'A'K\Lambda hL_V^{(i)}$  (where  $0 < \lambda < X^{1/24}$ ), and  $hL_V^{(i)}$  lies in a fixed compact set, we see that  $a = O(X^{1/6})$ ,  $b = O(X^{1/6})$ ,  $c = O(X^{1/6})$ ,  $ad = O(X^{2/6})$ ,  $bd = O(X^{2/6})$ , and  $ae = O(X^{2/6})$ . In particular, the latter estimates clearly imply that the number of points in  $\mathcal{R}_X(hL_V^{(i)})$  with  $a \neq 0$  and  $e = 0$  is  $O(X^{4/6+\epsilon})$ .

Let us now assume that  $a \neq 0$  and  $e \neq 0$ . We first estimate the number of forms that have a rational linear factor. The above estimates show that the number of possibilities for the quadruple  $(a, b, d, e)$  is at most  $O(X^{4/6+\epsilon})$ . If  $px + qy$  is a linear factor of  $f(x, y)$ , where  $p, q \in \mathbb{Z}$  are relatively prime, then  $p$  must be a factor of  $a$ , while  $q$  must be a factor of  $e$ ; they are thus both determined up to  $O(X^\epsilon)$  possibilities. Once  $p$  and  $q$  are determined, computing  $f(-q, p)$  and setting it equal to zero then uniquely determines  $c$  (if it is an integer at all) in terms of  $a, b, d, e, p, q$ . Thus the total number of forms  $f \in \mathcal{R}_X(hL_V^{(i)})$  having a rational linear factor and  $a \neq 0$  is  $O(X^{4/6+\epsilon})$ .

We now estimate the number of binary quartic forms in  $\mathcal{R}_X(hL_V^{(i)})$  that factor into two irreducible binary quadratic forms over  $\mathbb{Z}$ , say

$$ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4 = (px^2 + qxy + ry^2)\left(\frac{a}{p}x^2 + sxy + \frac{e}{r}y^2\right)$$

where  $p, q, r, s \in \mathbb{Z}$  and  $p, q, r$  are pairwise relatively prime. Since  $ae = O(X^{2/6})$  and  $a, e \neq 0$ , the number of possibilities for the pair  $(a, e)$  is  $O(X^{2/6+\epsilon})$ . We then see that  $p$  divides  $a$  and  $r$  divides  $e$ , and hence the number of possibilities for  $(p, r)$ , once  $a$  and  $e$  have been fixed, is bounded by  $O(X^\epsilon)$ .

Next, equating coefficients, we see that:

$$\begin{aligned} \frac{a}{p}q + ps &= b, \\ \frac{e}{r}q + rs &= d. \end{aligned} \tag{7}$$

We split into two cases. We first consider the case where  $\frac{ar}{pe} \neq \frac{p}{r}$ , i.e., the linear system (7) in the variables  $q$  and  $s$  is nonsingular. Then the values of  $b$  and  $d$  uniquely determine  $q$  and  $s$ , and so the total number of quadruples  $(a, b, d, e)$ —and hence the total number of octuples  $(a, b, d, e, p, r, q, s)$ —is at most  $O(X^{4/6+\epsilon})$ . Furthermore, once this octuple has been fixed, this also then determines  $c$  by equating coefficients of  $x^2y^2$ . Hence there are at most  $O(X^{4/6+\epsilon})$  possibilities for  $(a, b, c, d, e)$  in this case.

Next, we consider the case where  $\frac{ar}{pe} = \frac{p}{r}$ , so that the system (7) is singular. In this case, the value of  $b$  determines the value of  $d$  uniquely, namely  $d = (r/p)b$ . We have already seen that there are  $O(X^{2/6+\epsilon})$  possibilities for the quadruple  $(a, e, p, r)$ . Since there are only  $O(X^{1/6})$  choices for each of  $b$  and  $c$ , and then  $d$  is determined by  $b$ , the total number of choices for  $(a, b, c, d, e)$  is again  $O(X^{4/6+\epsilon})$ , as desired.  $\square$

Finally, we have the following lemma which bounds the number of  $\mathrm{GL}_2(\mathbb{Z})$ -equivalence classes of integral binary quartic forms having large stabilizers inside  $\mathrm{GL}_2(\mathbb{Z})$ ; we defer the proof to §3.8.

**Lemma 2.4** *Let  $h \in G_0$  be any element, where  $G_0$  is any fixed compact subset of  $\mathrm{GL}_2(\mathbb{R})$ . Then the number of integral binary quartic forms  $ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4 \in \mathcal{R}_X(hL_V^{(i)})$  whose stabilizer in  $\mathrm{GL}_2(\mathbb{Z})$  has size greater than 2 is  $O(X^{3/4+\epsilon})$ .*

### 2.3 Averaging and cutting off the cusp

Let  $G_0$  be a compact left  $K$ -invariant set in  $G$  that is the closure of a nonempty open set and in which every element has determinant greater than or equal to 1. Then for  $i = 0, 1, 2+$ , and  $2-$ , we may write

$$N(V_{\mathbb{Z}}^{(i)}; X) = \frac{\int_{h \in G_0} \#\{x \in \mathcal{F}hL \cap V_{\mathbb{Z}}^{\text{irr}} : H(x) < X\} dh}{n_i \cdot \int_{h \in G_0} dh}, \quad (8)$$

where  $V_{\mathbb{Z}}^{\text{irr}}$  denotes the set of irreducible elements in  $V_{\mathbb{Z}}^{\text{irr}}$  and  $L$  is equal to  $L_V^{(i)}$ . The denominator of the latter expression is an absolute constant  $C_{G_0}^{(i)}$  greater than zero.

More generally, for any  $\text{GL}_2(\mathbb{Z})$ -invariant subset  $S \subset V_{\mathbb{Z}}^{(i)}$ , let  $N(S; X)$  denote the number of irreducible  $\text{GL}_2(\mathbb{Z})$ -orbits in  $S$  having height less than  $X$ . Let  $S^{\text{irr}}$  denote the subset of irreducible points of  $S$ . Then  $N(S; X)$  can be similarly expressed as

$$N(S; X) = \frac{\int_{h \in G_0} \#\{x \in \mathcal{F}hL \cap S^{\text{irr}} : H(x) < X\} dh}{C_{G_0}^{(i)}}. \quad (9)$$

Given  $x \in V_{\mathbb{R}}^{(i)}$ , let  $x_L$  denote the unique point in  $L$  that is  $\text{GL}_2(\mathbb{R})$ -equivalent to  $x$ . We have

$$N(S; X) = \frac{1}{C_{G_0}^{(i)}} \sum_{\substack{x \in S^{\text{irr}} \\ H(x) < X}} \int_{h \in G_0} \#\{g \in \mathcal{F} : x = ghx_L\} dh. \quad (10)$$

For a given  $x \in S^{\text{irr}}$ , there exist a finite number of elements  $g_1, \dots, g_n \in \text{GL}_2(\mathbb{R})$  satisfying  $g_j x_L = x$ . We then have

$$\int_{h \in G_0} \#\{g \in \mathcal{F} : x = ghx_L\} dh = \sum_j \int_{h \in G_0} \#\{g \in \mathcal{F} : gh = g_j\} dh = \sum_j \int_{h \in G_0 \cap \mathcal{F}^{-1}g_j} dh.$$

As  $dh$  is an invariant measure on  $G$ , we have

$$\sum_j \int_{h \in G_0 \cap \mathcal{F}^{-1}g_j} dh = \sum_j \int_{g \in G_0 g_j^{-1} \cap \mathcal{F}^{-1}} dg = \sum_j \int_{g \in \mathcal{F}} \#\{h \in G_0 : gh = g_j\} dg = \int_{g \in \mathcal{F}} \#\{h \in G_0 : x = ghx_L\} dg.$$

Therefore,

$$N(S; X) = \frac{1}{C_{G_0}^{(i)}} \sum_{\substack{x \in S^{\text{irr}} \\ H(x) < X}} \int_{g \in \mathcal{F}} \#\{h \in G_0 : x = ghx_L\} dg. \quad (11)$$

$$= \frac{1}{C_{G_0}^{(i)}} \int_{g \in \mathcal{F}} \#\{x \in S^{\text{irr}} \cap gG_0L : H(x) < X\} dg \quad (12)$$

$$= \frac{1}{C_{G_0}^{(i)}} \int_{g \in N'(t)A' \Lambda K} \#\{x \in S^{\text{irr}} \cap n \begin{pmatrix} t^{-1} & \\ & t \end{pmatrix} \lambda k G_0 L : H(x) < X\} t^{-2} dn d^\times t d^\times \lambda dk. \quad (13)$$

Let us write  $B(n, t, \lambda, X) = n \begin{pmatrix} t^{-1} & \\ & t \end{pmatrix} \lambda G_0 L \cap \{x \in V^{(i)} : H(x) < X\}$ . As  $KG_0 = G_0$  and  $\int_K dk = 1$ , we have

$$N(S; X) = \frac{1}{C_{G_0}^{(i)}} \int_{g \in N'(t)A' \Lambda} \#\{x \in S^{\text{irr}} \cap B(n, t, \lambda, X)\} t^{-2} dn d^\times t d^\times \lambda. \quad (14)$$

To estimate the number of lattice points in  $B(n, t, \lambda, X)$ , we have the following proposition due to Davenport [22].

**Proposition 2.5** *Let  $\mathcal{R}$  be a bounded, semi-algebraic multiset in  $\mathbb{R}^n$  having maximum multiplicity  $m$ , and that is defined by at most  $k$  polynomial inequalities each having degree at most  $\ell$ . Let  $\mathcal{R}'$  denote the image of  $\mathcal{R}$  under any (upper or lower) triangular, unipotent transformation of  $\mathbb{R}^n$ . Then the number of integer lattice points (counted with multiplicity) contained in the region  $\mathcal{R}'$  is*

$$\text{Vol}(\mathcal{R}) + O(\max\{\text{Vol}(\bar{\mathcal{R}}), 1\}),$$

where  $\text{Vol}(\bar{\mathcal{R}})$  denotes the greatest  $d$ -dimensional volume of any projection of  $\mathcal{R}$  onto a coordinate subspace obtained by equating  $n-d$  coordinates to zero, where  $d$  takes all values from 1 to  $n-1$ . The implied constant in the second summand depends only on  $n$ ,  $m$ ,  $k$ , and  $\ell$ .

Although Davenport states the above lemma only for compact semi-algebraic sets  $\mathcal{R} \subset \mathbb{R}^n$ , his proof adapts without significant change to the more general case of a bounded semi-algebraic multiset  $\mathcal{R} \subset \mathbb{R}^n$ , with the same estimate applying also to any image  $\mathcal{R}'$  of  $\mathcal{R}$  under a unipotent triangular transformation.

By our construction of the  $L_V^{(i)}$ , the coefficients of the binary quartic forms in  $G_0L$  are all uniformly bounded. Let  $C$  be a constant that bounds the absolute value of the leading coefficient  $a$  of all the forms  $ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4$  in  $G_0L$ .

We then have the following lemma on the number of irreducible lattice points in  $B(n, t, \lambda, X)$ :

**Proposition 2.6** *The number of lattice points  $(a, b, c, d, e)$  in  $B(n, t, \lambda, X)$  with  $a \neq 0$  is*

$$\begin{cases} 0 & \text{if } \frac{C\lambda}{t} < 1; \\ \text{Vol}(B(n, t, \lambda, X)) + O(\max\{C^4 t^4 \lambda^{16}, 1\}) & \text{otherwise.} \end{cases}$$

**Proof:** If  $C\lambda/t < 1$ , then  $a = 0$  is the only possibility for an integral binary quartic form  $ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4$  in  $B(n, t, \lambda, X)$ , and any such binary cubic form is reducible. If  $C\lambda/t \geq 1$ , then  $\lambda$  and  $t$  are positive numbers bounded from below by  $(\sqrt{3}/2)/C$  and  $\sqrt{3}/2$  respectively. In this case, one sees that the projection of  $B(n, t, \lambda, X)$  onto  $a = 0$  has volume  $O(C^4 t^4 \lambda^{16})$ , while all other projections are also bounded by a constant times this. The lemma now follows from Proposition 2.5.  $\square$

In (14), since  $L$  (and therefore also  $G_0L$ ) only contains points with height at least 1, we observe (by the definition of  $B(n, t, \lambda, X)$ ) that the integrand will be nonzero only if  $t < C\lambda$  and  $\lambda < X^{1/24}$ . Thus we may write, up to an error of  $O(X^{2/3+\epsilon})$  due to the estimates on reducible forms in Lemma 2.3, that

$$N(V_{\mathbb{Z}}^{(i)}; X) = \frac{1}{C_{G_0}^{(i)}} \int_{\lambda=\sqrt{3}/(2C)}^{X^{1/24}} \int_{t=\sqrt{3}/2}^{C\lambda} \int_{N'(t)} (\text{Vol}(B(n, t, \lambda, X)) + O(\max\{C^4 t^4 \lambda^{16}, 1\})) t^{-2} dnd^\times td^\times \lambda. \quad (15)$$

The integral of the second summand is immediately evaluated to be  $O(X^{3/4})$ . Meanwhile, the integral of the first summand is

$$\frac{1}{C_{G_0}^{(i)}} \int_{h \in G_0} \text{Vol}(\mathcal{R}_X(hL)) dh - O\left(\int_{\lambda=\sqrt{3}/(2C)}^{X^{1/24}} \int_{t=C\lambda}^{\infty} \int_{N'(t)} \text{Vol}(B(n, t, \lambda, X)) t^{-2} dnd^\times td^\times \lambda\right). \quad (16)$$

However,  $\text{Vol}(\mathcal{R}_X(hL))$  is independent of  $h$ ; also, since  $\text{Vol}(B(n, t, \lambda, X)) = O(\lambda^{20})$ , by carrying out the integration in the second term of (16), we see that that this term is also  $O(X^{3/4})$ . We conclude that

$$N(V_{\mathbb{Z}}^{(i)}; X) = \text{Vol}(\mathcal{R}_X(L))/n_i + O(X^{3/4+\epsilon}). \quad (17)$$

To complete the proof of Theorem 2.1, it thus remains only to compute the volume  $\text{Vol}(\mathcal{R}_X(L))$ .

## 2.4 Computing the volume

Let  $i$  be equal to 0, 1, 2+, or 2-. In this section, we compute the volume of  $\mathcal{R}_X(L_V^{(i)})$ . To accomplish this, we define the usual subgroups  $\bar{N}, A_+, N$ , and  $\Lambda$  of  $\mathrm{GL}_2(\mathbb{R})$  as follows:

$$\begin{aligned}\bar{N} &= \{\bar{n}(\bar{u}) : \bar{u} \in \mathbb{R}\}, \text{ where } \bar{n}(\bar{u}) = \begin{pmatrix} 1 & \bar{u} \\ & 1 \end{pmatrix}, \\ A &= \{\alpha(t) : t \in \mathbb{R}^\times\}, \text{ where } \alpha(t) = \begin{pmatrix} t^{-1} & \\ & t \end{pmatrix}, \\ N &= \{n(u) : u \in \mathbb{R}\}, \text{ where } n(u) = \begin{pmatrix} 1 & \\ u & 1 \end{pmatrix}, \\ \Lambda &= \left\{ \begin{pmatrix} \lambda & \\ & \lambda \end{pmatrix} \right\} \text{ where } \lambda > 0.\end{aligned}$$

It is well-known that the product map  $\bar{N} \times A \times N \rightarrow \mathrm{SL}_2(\mathbb{R})$  maps bijectively onto a full measure set in  $\mathrm{SL}_2(\mathbb{R})$ . This decomposition gives a Haar measure  $dg$  on  $\mathrm{SL}_2(\mathbb{R})$ , namely,  $dg = t^{-2} d\bar{n} d^\times t dn = t^{-2} d\bar{u} d^\times t du$ .

Let  $R_V^{(i)} := \Lambda L_V^{(i)}$ . Then for each  $(I, J) \in \mathbb{R} \times \mathbb{R}$ , the set  $R_V^{(i)}$  contains at most one point  $p_{I,J}^{(i)}$  having invariants  $I$  and  $J$ . Let  $R_V^{(i)}(X)$  denote the set of all those points in  $R_V^{(i)}$  having height less than  $X$ . Then  $\mathrm{Vol}(\mathcal{R}_X(L_V^{(i)})) = \mathrm{Vol}(\mathcal{F}_{\mathrm{SL}_2} \cdot R_V^{(i)}(X))$ , where  $\mathcal{F}_{\mathrm{SL}_2}$  is the fundamental domain  $N'A'K$  for the action of  $\mathrm{SL}_2(\mathbb{Z})$  on  $\mathrm{SL}_2(\mathbb{R})$  (here  $N'$ ,  $A'$ , and  $K$  are as in (6)).

The set  $R_V^{(i)}$  is in canonical one-to-one correspondence with the set  $\{(I, J) \in \mathbb{R} \times \mathbb{R} : I^3 - J^2/4 > 0\}$  if  $i = 0, 2+$ , or  $2-$ , and with  $\{(I, J) \in \mathbb{R} \times \mathbb{R} : I^3 - J^2/4 < 0\}$  if  $i = 1$ . There is thus a natural measure on each of these sets  $R_V^{(i)}$ , given by  $dr = dI dJ$ . We have the following proposition:

**Proposition 2.7** *For any measurable function  $\phi$  on  $V_{\mathbb{R}}$ , we have*

$$\frac{2}{27n_i} \int_{R_V^{(i)}} \int_{\mathrm{SL}_2(\mathbb{R})} \phi(g \cdot p_{I,J}^{(i)}) dg dIdJ = \int_{V_{\mathbb{R}}^{(i)}} \phi(v) dv. \quad (18)$$

We know that  $\mathrm{SL}_2(\mathbb{R}) \cdot R_V^{(i)}$  is an  $n_i$ -fold cover of  $V_{\mathbb{R}}^{(i)}$ . The proposition then follows from a Jacobian computation and can be verified directly.

There is an alternative method of arriving at (18) that does not require one to use the explicit formulas for the fundamental sets  $L_V^{(i)}$ . We prove the following proposition, which will also be of use to us in the sequel.

**Proposition 2.8** *Let  $dg$  be the Haar measure on  $\mathrm{SL}_2(\mathbb{C})$  obtained from the usual  $\bar{N}AN$  decomposition. Let  $dv$  be the standard Euclidean measure on  $V_{\mathbb{C}}$ , the  $\mathbb{C}$ -vector space of all binary quartic forms with complex coefficients. Let  $R$  be a subset of  $V_{\mathbb{C}}$  that contains exactly one point  $p_{I,J}$  having invariants equal to  $I$  and  $J$  for each pair  $(I, J) \in \mathbb{C} \times \mathbb{C}$  and such that  $G_0 R$  is measurable for any measurable set  $G_0 \subset \mathrm{SL}_2(\mathbb{C})$ .*

*Then for any measurable function  $\phi$  on  $V_{\mathbb{C}}$ , we have*

$$\frac{2}{27n} \int_R \int_{\mathrm{SL}_2(\mathbb{C})} \phi(g \cdot p_{I,J}) dg dIdJ = \int_{v \in V_{\mathbb{C}}} \phi(v) dv$$

where  $n$  is equal to 8, the size of the stabilizer of any point  $v \in V_{\mathbb{C}}$  having nonzero discriminant.

**Proof:** We first prove that  $\mathrm{SL}_2(\mathbb{C}) \cdot R$  is an 8-fold cover of a full measure set in  $V_{\mathbb{C}}$ . To do this, we consider the set  $R_0$  consisting of the forms  $q_{I,J} = x^3y - \frac{I}{3}xy^3 - \frac{J}{27}y^4$  for each  $(I, J) \in \mathbb{C} \times \mathbb{C}$ . Note that  $q_{I,J}$  has invariants equal to  $I$  and  $J$ .

Now if  $p_{I,J} \in R$  has nonzero invariants  $I$  and  $J$  such that  $4I^3 - J^2 \neq 0$ , then there exist 8 elements  $g_{I,J} \in \mathrm{SL}_2(\mathbb{C})$  such that  $g_{I,J} \cdot p_{I,J} = q_{I,J}$ . This can be seen in the following steps. We first apply an  $\mathrm{SL}_2(\mathbb{C})$ -transformation on  $p_{I,J}$  to send one of its four distinct roots in  $\mathbb{P}^1(\mathbb{C})$  to the point  $[1 : 0]$ . The resulting quartic is of the form  $bx^3y + cx^2y^2 + dxy^3 + ey^4$  with  $b \neq 0$ . Secondly, we use a diagonal matrix and then

a lower triangular matrix of  $\mathrm{SL}_2(\mathbb{C})$  to transform the quartic into the form  $x^3y + d'xy^3 + e'y^4$ . Because this quartic still has invariants equal to  $I$  and  $J$ , we see that it is in fact equal to  $q_{I,J}$ . Finally, identically as in the proof of Lemma 2.2, the stabilizer in  $\mathrm{SL}_2(\mathbb{C})$  of any binary quartic form in  $V_{\mathbb{C}}$  having invariants  $I \neq 0$  and  $J \neq 0$  has size 8. Thus  $\mathrm{SL}_2(\mathbb{C}) \cdot R$  is an  $n$ -fold cover of a full measure set in  $V_{\mathbb{C}}$ .

The proposition can be checked directly for  $R = R_0$  via a Jacobian computation. We may now use this fact to derive the result for general  $R$  as follows:

$$\begin{aligned} \int_{v \in \mathrm{SL}_2(\mathbb{C}) \cdot R} \phi(v) dv &= \frac{2}{27} \int_{R_0} \int_{\mathrm{SL}_2(\mathbb{C})} \phi(g \cdot q_{I,J}) dg dr = \frac{2}{27} \int_R \int_{\mathrm{SL}_2(\mathbb{C})} \phi(gg_{I,J} \cdot p_{I,J}) dg dr \\ &= \frac{2}{27} \int_R \int_{\mathrm{SL}_2(\mathbb{C})} \phi(g \cdot p_{I,J}) dg dr, \end{aligned}$$

where the last equality follows from the fact that  $\mathrm{SL}_2(\mathbb{C})$  is a *unimodular* group (see [34, Chapter 8]), i.e., the left Haar measure  $dg$  is also a right Haar measure on  $\mathrm{SL}_2(\mathbb{C})$ .  $\square$

Proposition 2.8, along with the principle of permanence of identities, implies an analogous result when  $\mathbb{C}$  is replaced with  $\mathbb{R}$  or  $\mathbb{Q}_p$  for any prime  $p$ . In particular, it implies Proposition 2.7, which may now be used to compute the volume of the multiset  $\mathcal{R}_X(L_V^{(i)})$ :

$$\int_{\mathcal{R}_X(L_V^{(i)})} dv = \int_{\mathcal{F}_{\mathrm{SL}_2} \cdot R_V^{(i)}(X)} dv = \frac{2}{27} \int_{R_V^{(i)}(X)} \int_{\mathcal{F}_{\mathrm{SL}_2}} dg dI dJ = \frac{2}{27} \cdot \zeta(2) \int_{R_V^{(i)}(X)} dI dJ. \quad (19)$$

When  $i = 0, 2+$ , or  $2-$ , we compute  $\int_{R_V^{(i)}(X)} dI dJ$  to be

$$\int_{I=0}^{X^{1/3}} \int_{J=-2I^{3/2}}^{2I^{3/2}} dI dJ = \int_{I=0}^{X^{1/3}} 4I^{3/2} dI = \frac{8}{5} I^{5/2} \Big|_0^{X^{1/3}} = \frac{8}{5} X^{5/6}. \quad (20)$$

Meanwhile,  $\int_{R_V^{(1)}(X)} dI dJ$  is equal to

$$\int_{I=-X^{1/3}}^{X^{1/3}} \int_{j=-2X^{1/2}}^{2X^{1/2}} dI dJ - \mathrm{Vol}(R_V^{(0)}(X)) = 8X^{5/6} - \frac{8}{5} X^{5/6} = \frac{32}{5} X^{5/6}. \quad (21)$$

We conclude that

$$\mathrm{Vol}(\mathcal{R}_X(L_V^{(i)})) = \begin{cases} \frac{16}{135} \cdot \zeta(2) X^{5/6} & \text{for } i = 0, 2+, \text{ and } 2-; \\ \frac{64}{135} \cdot \zeta(2) X^{5/6} & \text{for } i = 1. \end{cases} \quad (22)$$

As  $n_0 = n_{2+} = n_{2-} = 4$  and  $n_1 = 2$ , equations (17) and (22) now immediately imply Theorem 2.1.

To deduce Theorem 1.8 from Theorem 2.1, we require a count of the number of eligible pairs  $(I, J) \in \mathbb{Z} \times \mathbb{Z}$  satisfying  $H(I, J) < X$ . We use the following lemma which we prove in §3.8:

**Lemma 2.9** *The set of eligible  $(I, J) \in \mathbb{Z} \times \mathbb{Z}$  is a union of 9 lattices, each of index 243 in  $\mathbb{Z} \times \mathbb{Z}$ .*

The next proposition is now a simple application of Proposition 2.5.

**Proposition 2.10** *Let  $N_{I,J}^+(X)$  and  $N_{I,J}^-(X)$  denote the number of eligible  $(I, J) \in \mathbb{Z} \times \mathbb{Z}$  satisfying  $H(I, J) < X$  that have positive discriminant and negative discriminant, respectively. Then we have*

- (a)  $N_{I,J}^+(X) = \frac{8}{135} X^{5/6} + O(X^{1/2});$
- (b)  $N_{I,J}^-(X) = \frac{32}{135} X^{5/6} + O(X^{1/2}).$



**Proof:** Let  $R_X^+$  denote the set  $\{(i, j) \in \mathbb{R}^2 : |i| < X^{1/3}, |j| < 2X^{1/2}, 4i^3 - j^2 > 0\}$  and let  $R_X^-$  denote the set  $\{(i, j) \in \mathbb{R}^2 : |i| < X^{1/3}, |j| < 2X^{1/2}, i^3 - j^2 < 0\}$ . The sizes of the projections of  $R_X^\pm$  onto smaller-dimensional coordinate hyperplanes are all bounded by  $O(X^{1/2})$ . Using Proposition 2.5 and Lemma 2.9 we then see that  $N_{I,J}^\pm(X) = \frac{9}{243} \text{Vol}(R_X^\pm) + O(X^{1/2})$ . The volumes of  $R_X^+$  and  $R_X^-$  were computed in (20) and (21), respectively, and the proposition follows.  $\square$

Theorem 1.8 now follows from Theorem 2.1 and Proposition 2.10.

## 2.5 Congruence conditions

In this subsection, we prove a version of Theorem 2.1 where we count integral binary quartic forms satisfying any finite specified set of congruence conditions.

Suppose  $S$  is a subset of  $V_{\mathbb{Z}}$  defined by finitely many congruence conditions. We may assume that  $S \subset V_{\mathbb{Z}}$  is defined by congruence conditions modulo some integer  $m$ . Then  $S$  may be viewed as the union of (say)  $k$  translates  $\mathcal{L}_1, \dots, \mathcal{L}_k$  of the lattice  $m \cdot V_{\mathbb{Z}}$ . For each such lattice translate  $\mathcal{L}_j$ , we may use formula (14) and the discussion following that formula to compute  $N(\mathcal{L}_j \cap V_{\mathbb{Z}}^{(i)}; X)$ , where each  $d$ -dimensional volume is scaled by a factor of  $1/m^d$  to reflect the fact that our new lattice has been scaled by a factor of  $m$ . With these scalings, the maximum volume of the projections of  $B(n, t, \lambda, X)$  is seen to be at most  $O(t^4 \lambda^{16})$ . Analogous to Proposition 2.6, we see that the number of points  $(a, b, c, d, e)$  in  $B(n, t, \lambda, X) \cap \mathcal{L}_j$  with  $a \neq 0$  is

$$\begin{cases} 0 & \text{if } \frac{C\lambda}{t} < 1; \\ \frac{1}{m^5} \text{Vol}(B(n, t, \lambda, X)) + O(\max\{C^4 t^4 \lambda^{16}, 1\}) & \text{otherwise.} \end{cases}$$

Carrying out the integral for  $N(\mathcal{L}_j \cap V_{\mathbb{Z}}^{(i)}; X)$  as in (17), we obtain, up to an error of  $O(X^{2/3+\epsilon})$  corresponding to the reducible points in Lemma 2.3, that

$$N(\mathcal{L}_j \cap V_{\mathbb{Z}}^{(i)}; X) = \frac{\text{Vol}(\mathcal{R}_X(L_V^{(i)}))}{n_i \cdot m^5} + O(X^{3/4+\epsilon}).$$

Summing over  $j$ , we thus obtain

$$N(S \cap V_{\mathbb{Z}}^{(i)}; X) = \frac{k \text{Vol}(\mathcal{R}_X(L_V^{(i)}))}{n_i \cdot m^5} + O(X^{3/4+\epsilon}). \quad (23)$$

For any set  $S$  in  $V_{\mathbb{Z}}$  that is definable by congruence conditions, let us denote by  $\mu_p(S)$  the  $p$ -adic density of the  $p$ -adic closure of  $S$  in  $V_{\mathbb{Z}_p}$ , where we normalize the additive measure  $\mu_p$  on  $V_{\mathbb{Z}_p}$  so that  $\mu_p(V_{\mathbb{Z}_p}) = 1$ . We then have the following theorem:

**Theorem 2.11** *Suppose  $S$  is a subset of  $V_{\mathbb{Z}}$  defined by finitely many congruence conditions. Then we have*

$$N(S \cap V_{\mathbb{Z}}^{(i)}; X) = N(V_{\mathbb{Z}}^{(i)}; X) \prod_p \mu_p(S) + O(X^{3/4+\epsilon}), \quad (24)$$

where  $\mu_p(S)$  denotes the  $p$ -adic density of  $S$  in  $V_{\mathbb{Z}}$ , and where the implied constant depends only on  $S$ .

Theorem 2.11 follows from Equations (17) and (23), together with the identity  $km^{-5} = \prod_p \mu_p(S)$ .

## 3 The mean number of 2-torsion elements in the class groups of monogenic cubic fields

In this section, we determine the mean size of the 2-torsion subgroup of the class group of *monogenic* cubic fields, where these cubic fields are ordered by the heights of their defining cubic polynomials. In particular, we prove Theorem 1.9 (assuming certain uniformity estimates which will be proven in Section 4).

### 3.1 The parametrization of monogenic cubic rings

We work over a principal ideal domain  $R$  that we will primarily assume to be equal to  $\mathbb{Z}$ ,  $\mathbb{Q}$ ,  $\mathbb{R}$ ,  $\mathbb{Z}_p$ , or  $\mathbb{Q}_p$  for some prime  $p$ .

Let  $U_R$  denote the space of *binary cubic forms* over  $R$ , i.e., polynomials of the form  $g(X, Y) = qX^3 + rX^2Y + sXY^2 + tY^3$  with  $q, r, s, t \in R$ . We call  $qX^3$  the *leading term* of  $g(X, Y)$ , and  $q$  the *leading coefficient*. A *monic binary cubic form* is defined to be a binary cubic form whose leading coefficient is equal to 1, i.e., it is of the form  $X^3 + rX^2Y + sXY^2 + tY^3$ . We denote the set of monic binary cubic forms over  $R$  by  $U_{R,1} \subset U_R$ .

Every monic binary cubic form  $g(X, Y) = X^3 + rX^2Y + sXY^2 + tY^3$  naturally gives rise to a cubic ring  $C = R[X]/(X^3 + rX^2 + sX + t)$  over  $R$  which is *monogenic*, i.e., the ring  $C$  is generated over  $R$  by a single element  $X$ . The element  $X$  is then said to be a *monogenizer* of  $C$ . We define a *monogenized* cubic ring to be a pair  $(C, x)$  such that  $C$  is a monogenic cubic ring over  $R$  and  $x$  is a monogenizer of  $C$ .

Two monogenized cubic rings  $(C, x)$  and  $(C', x')$  are said to be *isomorphic* if there exists a ring isomorphism  $\phi : C \rightarrow C'$  such that  $\phi(x) = x' + u$  for some  $u \in R$ . In particular, the monogenized cubic rings  $(C, x)$  and  $(C, x + u)$  are isomorphic. Thus it is natural to consider the corresponding monic integral binary cubic forms  $g(X, Y)$  and  $g(X + u, Y)$  to be equivalent too, leading to a natural action of the subgroup

$$F_{R,1} = \left\{ \begin{pmatrix} 1 & \\ u & 1 \end{pmatrix} : u \in R \right\} \quad (25)$$

of  $\mathrm{SL}_2(R)$  on  $U_{R,1}$ , given by  $\gamma \cdot g(X, Y) = g((X, Y) \cdot \gamma)$ .

We may summarize the above discussion in the following proposition, which is essentially the same as Proposition 4.3.1 in [41].

**Proposition 3.1** *The  $F_{R,1}$ -orbits on  $U_{R,1}$  are in bijective correspondence with isomorphism classes of monogenized cubic rings over  $R$ .*

Next, if  $g(X, Y) = X^3 + rX^2Y + sXY^2 + tY^3$ , then one easily sees that the quantities

$$\begin{aligned} I(g) &:= r^2 - 3s, \\ J(g) &:= -2r^3 + 9rs - 27t \end{aligned} \quad (26)$$

are invariant under the action of  $F_{R,1}$ . The discriminant  $\Delta(g)$  of the binary cubic form  $g$  can be expressed in terms of these basic invariants  $I(g)$  and  $J(g)$ , namely,  $\Delta(g) = (4I^3(g) - J^2(g))/27$ . We again define the *height* of  $g$  by

$$H(g) := H(I, J) = \max\{|I^3(g)|, J^2(g)/4\}.$$

If  $C = R[X]/g(X)$ , then the *height*  $H(C, X)$  of the monogenized cubic ring  $(C, X)$  is naturally defined by  $H(C, X) := H(g)$ .

### 3.2 Counting monogenized cubic rings by height

We begin with the observation that the  $F_{\mathbb{Z},1}$ -orbits on  $U_{\mathbb{Z},1}$  are specified uniquely by their invariants  $(I, J)$ . Indeed, the set

$$\mathcal{S} = \{X^3 + rX^2Y + sXY^2 + tY^3 : r \in \{-1, 0, 1\}; s, t \in \mathbb{Z}\} \quad (27)$$

is clearly a fundamental set for the action of  $F_{\mathbb{Z},1}$  on  $U_{\mathbb{Z},1}$ . Given  $g(X, Y) = X^3 + rX^2Y + sXY^2 + tY^3 \in \mathcal{S}$  with invariants  $I, J$ , the value of  $r \in \{-1, 0, 1\}$  is determined by the value of  $J \pmod{3}$  since  $J \equiv -2r^3 \pmod{3}$ . Once  $r$  is so determined,  $s$  is determined by the value of  $I$  and then  $t$  is determined by the value of  $J$ .

This raises the question as to which pairs  $(I, J)$  of invariants can actually arise for a monic binary cubic form. This question is answered by the following proposition:

**Proposition 3.2** *The pair  $(I, J)$  occurs as the invariants of a monic binary cubic form if and only if one of the following conditions is satisfied:*

- (a)  $I \equiv 0 \pmod{3}$  and  $J \equiv 0 \pmod{27}$ ;
- (b)  $I \equiv 1 \pmod{9}$  and  $J \equiv \pm 2 \pmod{27}$ ;
- (c)  $I \equiv 4 \pmod{9}$  and  $J \equiv \pm 16 \pmod{27}$ ;
- (d)  $I \equiv 7 \pmod{9}$  and  $J \equiv \pm 7 \pmod{27}$ .

**Proof:** Recall that  $\mathcal{S}$ , as defined in (27), is a fundamental set for the action of  $F_{\mathbb{Z},1}$  on  $U_{\mathbb{Z},1}$ . Therefore, a pair  $(I, J)$  occurs as the invariants of some monogenic binary cubic form in  $U_{\mathbb{Z},1}$  if and only if it occurs as the invariants of some element in  $\mathcal{S}$ .

Suppose that  $I, J$  give the invariants of  $g(X, Y) = X^3 + rX^2Y + sXY^2 + tY^3 \in \mathcal{S}$ , i.e.,  $I = r^2 - 3s$  and  $J = -2r^3 + 9rs - 27t$ . If  $I = r^2 - 3s \equiv 0 \pmod{3}$  then  $r = 0$ , implying that  $27 \mid J$ . This is condition (a).

If  $I$  is not divisible by 3, then  $r = 1$  or  $-1$  and thus  $I \equiv 1 \pmod{3}$ . Thus  $I$  must be congruent to 1, 4, or 7 (mod 9), which happens exactly when  $s$  is congruent to 0, 2, or 1 (mod 3), respectively. Because  $r^2 = 1$ , we see that  $J \equiv r(9s - 2) \pmod{27}$ . It follows that  $I \equiv 1, 4, 7 \pmod{9}$  corresponds to  $J \equiv \pm 2, \pm 16, \pm 7 \pmod{27}$ , respectively, yielding conditions (b), (c), and (d).

The converse also follows easily by reversing the above arguments.  $\square$

We now have the following lemma which bounds the number of reducible monic binary cubic forms, up to  $F_{\mathbb{Z},1}$ -equivalence, having height less than  $X$ .

**Lemma 3.3** *Let  $\mathcal{S}$  be the fundamental set for the action of  $F_{\mathbb{Z},1}$  on  $U_{\mathbb{Z},1}$  as defined by (27). Then the number of forms  $g \in \mathcal{S}$  such that  $g$  is reducible and  $H(g) < X$  is  $O(X^{1/2+\epsilon})$ .*

**Proof:** First, we note that if  $g(x, y) = x^3 + rx^2y + sxy^2 + ty^3 \in \mathcal{S}$  satisfies  $H(g) < X$ , then since  $|I(g)|^3 = |r^2 - 3s|^3 \leq H(g) < X$ , we see that  $s = O(X^{1/3})$ . Then since  $J(g)^2/4 = (2r^3 + 9rs - 27t)^2/4 \leq H(g) < X$ , this in turn implies that  $t = O(X^{1/2})$ .

Let us now count such forms  $g$  that are reducible. If  $g(x, y) = x^3 + rx^2y + sxy^2 + ty^3 \in \mathcal{S}$  satisfies  $t = 0$ , then  $g$  is reducible, and the number of forms  $g \in \mathcal{S}$  with  $t = 0$  and  $H(g) < X$  is the number of possible values for  $r$  and  $s$ , namely  $3 \cdot O(X^{1/3}) = O(X^{1/3})$ .

Next, we consider those reducible forms  $g(x, y) = x^3 + rx^2y + sxy^2 + ty^3 \in \mathcal{S}$  satisfying  $H(g) < X$  and  $t \neq 0$ . If  $x - my$  is a factor of  $g$ , then  $m \mid t$ . Therefore, if we fix  $t \neq 0$ , then there are at most  $t^\epsilon = O(X^\epsilon)$  choices for  $m$ . Moreover, once  $r, t$ , and  $m$  are fixed, then setting  $g(m, 1)$  equal to 0 determines  $s$ . Since  $t = O(X^{1/2})$ , and there are at most 3 possible values for  $r$ , it follows that there are at most  $O(X^{1/2+\epsilon})$  such reducible forms  $g \in \mathcal{S}$  with height less than  $X$ . This completes the proof.  $\square$

Let  $U_{\mathbb{Z},1}^{(0)}$  (resp.  $U_{\mathbb{Z},1}^{(1)}$ ) denote the set of monic binary cubic forms over  $\mathbb{Z}$  having positive (resp. negative) discriminant. For an  $F_{\mathbb{Z},1}$ -invariant set  $S \subset U_{\mathbb{Z},1}$ , let  $N(S; X)$  denote the number of  $F_{\mathbb{Z},1}$ -equivalence classes of irreducible elements  $g \in S$  satisfying  $H(g) < X$ . The following proposition gives asymptotic formulas for  $N(U_{\mathbb{Z},1}^{(0)}; X)$  and  $N(U_{\mathbb{Z},1}^{(1)}; X)$ :

**Proposition 3.4** *We have*

- (a)  $N(U_{\mathbb{Z},1}^{(0)}; X) = \frac{8}{135}X^{5/6} + O(X^{1/2+\epsilon});$
- (b)  $N(U_{\mathbb{Z},1}^{(1)}; X) = \frac{32}{135}X^{5/6} + O(X^{1/2+\epsilon}).$

**Proof:** A monic binary cubic form is determined up to equivalence by its invariants  $I$  and  $J$ . It therefore suffices to count integer pairs  $(I, J)$  such that:  $|I| < X^{1/3}$ ,  $|J| < 2X^{1/2}$ , and  $(I, J)$  occurs as the invariants of some monic binary cubic form  $g$ . Proposition 3.2 now reduces the proof of Proposition 3.4 to counting integer points  $\{(I, J) : H(I, J) < X\}$  satisfying the congruences (a)–(d) of Proposition 3.2. The proof of Proposition 2.10 now yields the result.  $\square$

We saw in §3.1 that isomorphism classes of monogenized cubic rings are parametrized by  $F_{\mathbb{Z},1}$ -orbits on  $U_{\mathbb{Z},1}$ . Since irreducible forms in  $U_{\mathbb{Z},1}$  correspond to monogenized orders in cubic fields, we have the following corollary:

**Corollary 3.5** *Let  $M_3^{(0)}(X)$  (resp.  $M_3^{(1)}(X)$ ) denote the number of isomorphism classes of monogenized orders in cubic fields having height less than  $X$  and positive (resp. negative) discriminant. Then*

$$(a) \quad M^{(0)}(X) = \frac{8}{135}X^{5/6} + O(X^{1/2+\epsilon});$$

$$(b) \quad M^{(1)}(X) = \frac{32}{135}X^{5/6} + O(X^{1/2+\epsilon}).$$

### 3.3 The parametrization of quartic rings having monogenic cubic resolvent rings

In this subsection, we discuss the parametrization of *quartic rings* over a base ring, i.e., rings that are free of rank 4 as a module over the base ring. For simplicity, we choose our base ring  $R$  to be a principal ideal domain having characteristic not equal to 2, which we will again primarily assume to be  $\mathbb{Z}$ ,  $\mathbb{Q}$ ,  $\mathbb{R}$ ,  $\mathbb{Z}_p$ , or  $\mathbb{Q}_p$  for some  $p$ .

Let  $W_R$  denote the space of pairs  $(A, B)$  of ternary quadratic forms having coefficients in  $R$ . We write an element  $(A, B) \in W_R$  as a pair of  $3 \times 3$  symmetric matrices via

$$2 \cdot (A, B) = \left( \begin{bmatrix} 2a_{11} & a_{12} & a_{13} \\ a_{12} & 2a_{22} & a_{23} \\ a_{13} & a_{23} & 2a_{33} \end{bmatrix}, \begin{bmatrix} 2b_{11} & b_{12} & b_{13} \\ b_{12} & 2b_{22} & b_{23} \\ b_{13} & b_{23} & 2b_{33} \end{bmatrix} \right),$$

where  $a_{ij}, b_{ij} \in R$ . The group  $G_R = \mathrm{GL}_2(R) \times \mathrm{SL}_3(R)$  acts naturally on the space  $V_R$ . Namely, an element  $g_3 \in \mathrm{SL}_3(R)$  acts on  $V_R$  by  $g_3 \cdot (A, B) = (g_3 A g_3^t, g_3 B g_3^t)$ , while an element  $g_2 = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \in \mathrm{GL}_2(R)$  acts by  $g_2 \cdot (A, B) = (pA + qB, rA + sB)$ . We have the following theorem parametrizing quartic algebras over  $R$ :

**Theorem 3.6 ([4])** *There is a canonical bijection between  $G_R$ -orbits on  $W_R$  and isomorphism classes of pairs  $(Q, C)$ , where  $Q$  is a quartic ring over  $R$  and  $C$  is a cubic resolvent ring of  $Q$ .*

A *cubic resolvent ring* of a quartic ring  $Q$  is a cubic ring  $C$  equipped with a certain quadratic resolvent mapping  $\phi_{4,3} : Q/R \rightarrow C/R$ ; see [4, §2.3, Definition 8] for more details. If  $(A, B) \in W_R$ , and  $(Q, C)$  is the corresponding pair of rings, then  $C$  can be explicitly described as follows. Let  $g(X, Y) = 4 \cdot \det(AX - BY)$ . Then  $g$  is a binary cubic form over  $R$ , and  $C$  is the cubic ring corresponding to the form  $g$  under the Delone–Faddeev bijection [26]. (A complete description of the construction of  $(Q, C)$  can be found in [4]).

Since  $G_R = \mathrm{GL}_2(R) \times \mathrm{SL}_3(R)$  acts on  $W_R$ , the groups  $\mathrm{GL}_2(R) \subset G_R$  and  $\mathrm{SL}_3(R) \subset G_R$  also act on  $W_R$  by restriction. Note that the action of  $\mathrm{SL}_3(R)$  fixes the binary cubic form  $g(X, Y) = 4 \cdot \det(AX - BY)$ , for any point  $(A, B)$  in  $W_R$ .

Now let  $V_R$  denote the space of binary quartic forms with coefficients in  $R$ . We consider the following twisted action of  $\mathrm{GL}_2(R)$  on  $V_R$ . Let  $\gamma \in \mathrm{GL}_2(R)$  and  $f \in V_R$ . Then

$$\gamma \cdot f(x, y) = \frac{1}{\det(\gamma)^2} f((x, y) \cdot \gamma). \quad (28)$$

Notice that for  $R = \mathbb{Z}$ , the action (28) is the same as the non-twisted action but for  $R = \mathbb{R}$  it is different. From now on, whenever we refer to an action of  $\mathrm{GL}_2(R)$  on  $V_R$  we shall be talking about the action (28). Note that the center of  $\mathrm{GL}_2(R)$  acts trivially on  $V_R$  under this action. Therefore the action descends to one of  $\mathrm{PGL}_2(R)$  on  $V_R$ , and it is clear that the quantities  $I$  and  $J$  remain invariant under this action.

The space of integral binary quartic forms embeds naturally into  $W_R$  via the map

$$\phi : ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4 \mapsto \left( \begin{bmatrix} & & 1/2 \\ & -1 & \\ 1/2 & & \end{bmatrix}, \begin{bmatrix} a & b/2 & 0 \\ b/2 & c & d/2 \\ 0 & d/2 & e \end{bmatrix} \right). \quad (29)$$

We denote the first matrix in the above pair by  $A_1$ .

Let  $f(x, y) = ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4 \in V_R$  be a binary quartic form and let us write  $\phi(f)$  as  $(A_1, B_f)$ . Let  $(Q, C)$  be the pair of rings associated to  $(A_1, B_f)$ . As we remarked earlier,  $C$  is the cubic ring

associated to the binary cubic form  $g(X, Y) = 4 \cdot \det(A_1 X - B_f Y)$ , and  $g$  is monic because  $\det(A_1) = 1/4$ . In fact, we compute  $g$  to be:

$$g(X, Y) = X^3 + cX^2Y + (bd - 4ae)XY^2 + (ad^2 + b^2e - 4ace)Y^3. \quad (30)$$

Moreover, we find that  $I(f) = I(g)$  and  $J(f) = J(g)$ . We define the *resolvent binary cubic form* of  $f$  to be  $g$ .

Note that the group  $F_{R,1} \subset \mathrm{GL}_2(R) \subset G_R$  acts on the subset  $W_{R,1} := \{(A, B) : A = A_1\}$  inside  $W_R$ , and every  $F_{R,1}$ -equivalence class of  $W_{R,1}$  contains a unique element  $(A_1, B)$  such that the top right entry of  $B$  is equal to 0. Therefore the space of binary quartic forms  $V_R$  maps to the set of  $F_{R,1}$ -equivalence classes on  $W_{R,1}$  via the composite map

$$V_R \rightarrow W_{R,1} \rightarrow F_{R,1} \backslash W_{R,1}.$$

From the representation of  $\mathrm{PGL}_2(R)$  on binary quadratic forms  $px^2 - 2qxy + ry^2$ , we obtain a map

$$\begin{aligned} \rho : \mathrm{PGL}_2(R) &\rightarrow \mathrm{SL}_3(R), \text{ given explicitly by} \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} &\mapsto \frac{1}{ad - bc} \begin{pmatrix} d^2 & cd & c^2 \\ 2bd & ad + bc & 2ac \\ b^2 & ab & a^2 \end{pmatrix}. \end{aligned} \quad (31)$$

The image of  $\mathrm{PGL}_2(R)$  is contained in the orthogonal group  $\mathrm{SO}(A_1, R)$  because  $A_1$  is the Gram matrix of the ternary form  $q^2 - pr$ . In fact, the image of  $\mathrm{PGL}_2(R)$  is equal to  $\mathrm{SO}(A_1, R)$ ; the proof parallels that of Lemma 4.4.2 in [41].

It is easily checked that the map  $\phi$  is equivariant with respect to  $\rho$ ; i.e.,  $\phi(\gamma \cdot f) = \rho(\gamma) \cdot \phi(f)$  for  $\gamma \in \mathrm{PGL}_2(R)$  and  $f \in V_R$ . Therefore, we have the following theorem:

**Theorem 3.7** *The map  $\phi$  defined by (29) gives a canonical bijection between  $\mathrm{PGL}_2(R)$ -classes on  $V_R$  and  $F_{R,1} \times \mathrm{SO}(A_1, R)$ -classes on  $W_{R,1}$ .*

We now restrict ourselves to the case where  $R = \mathbb{Z}$ . It is known that  $A_1$  is the Gram matrix of the unique quadratic form over  $\mathbb{Z}$  up to  $\mathrm{SL}_3(\mathbb{Z})$ -equivalence having determinant equal to  $1/4$ . Therefore, every pair  $(A, B) \in V_{\mathbb{Z}}$  such that  $\det(A) = 1/4$  is  $G_{\mathbb{Z}}$ -equivalent to  $(A_1, B')$  for some integral ternary quadratic form  $B'$ . We thus have the following corollary to Theorem 3.7, which is Theorem 4.1.1 in [41].

**Corollary 3.8 (Wood [41])** *There is a canonical bijection between  $\mathrm{GL}_2(\mathbb{Z})$ -equivalence classes of integral binary quartic forms and isomorphism classes of pairs  $(Q, C)$ , where  $Q$  is a quartic ring over  $\mathbb{Z}$  and  $C$  is a monogenized cubic resolvent ring of  $Q$ .*

In §3.8, we will also have occasion to use Theorem 3.7 when  $R = \mathbb{R}$ .

Combining Theorem 2.1 and Corollary 3.8, we obtain the following theorem.

**Theorem 3.9** *Let  $M_4^{(i)}(X)$  denote the number of isomorphism classes of pairs  $(Q, C)$  where  $Q$  is an order in a quartic field having  $4 - 2i$  real embeddings and  $C$  is a monogenized cubic resolvent ring of  $Q$  having height less than  $X$ . Then*

$$\begin{aligned} \text{(a)} \quad M_4^{(0)}(X) &= \frac{4}{135} \zeta(2) X^{5/6} + O(X^{3/4+\epsilon}); \\ \text{(b)} \quad M_4^{(1)}(X) &= \frac{32}{135} \zeta(2) X^{5/6} + O(X^{3/4+\epsilon}); \\ \text{(a)} \quad M_4^{(2)}(X) &= \frac{8}{135} \zeta(2) X^{5/6} + O(X^{3/4+\epsilon}). \end{aligned}$$

### 3.4 The parametrization of 2-torsion elements in the class groups of monogenic cubic fields

A cubic ring  $C$  over  $\mathbb{Z}$  is said to be *maximal* if it is not strictly contained in any other cubic ring over  $\mathbb{Z}$ . If a cubic ring  $C$  is maximal, denote by  $\text{Cl}_2(C)$  and  $\text{Cl}_2^+(C)$  the 2-torsion subgroup of the class group and of the narrow class group of  $C$ , respectively. Denote the abelian groups dual to  $\text{Cl}_2(C)$  and  $\text{Cl}_2^+(C)$  by  $\text{Cl}_2(C)^*$  and  $\text{Cl}_2^+(C)^*$ , respectively.

If we restrict Corollary 3.8 to maximal orders, then we obtain the following theorem:

**Theorem 3.10** *Let  $g$  be an irreducible monic integral binary cubic form having invariants  $I$  and  $J$  such that the cubic ring  $C(g) := \mathbb{Z}[x]/g(x)$  is maximal.*

- (a) *Suppose  $4I^3 - J^2 > 0$ . Then there is a canonical bijection between elements of  $\text{Cl}_2^+(C)^*$  and  $\text{GL}_2(\mathbb{Z})$ -equivalence classes of integral binary quartic forms having invariants  $I$  and  $J$ .*

*Under this bijection, elements of  $\text{Cl}_2(C)^* \subset \text{Cl}_2^+(C)^*$  correspond to those binary quartic forms having 4 real roots.*

- (b) *Suppose  $4I^3 - J^2 < 0$ . Then there is a canonical bijection between elements of  $\text{Cl}_2(C)^* = \text{Cl}_2^+(C)^*$  and  $\text{GL}_2(\mathbb{Z})$ -equivalence classes of integral binary quartic forms having invariants  $I$  and  $J$ .*

*Moreover, the set of binary quartic forms having an integer linear factor and invariants equal to  $I$  and  $J$  all lie in the same  $\text{GL}_2(\mathbb{Z})$ -equivalence class; this equivalence class corresponds to the identity element of  $\text{Cl}_2^+(C)^*$ .*

**Proof:** If  $Q$  is a maximal quartic order, then we say that  $Q$  is *overramified* at the prime  $p \in \mathbb{Z}$  if  $p$  factors into primes in  $Q \otimes_{\mathbb{Z}} \mathbb{Z}_p$  as  $P^4$ ,  $P^2$ , or  $P_1^2 P_2^2$ ; similarly,  $Q$  is *overramified* at the archimedean prime of  $\mathbb{Z}$  if  $Q \otimes \mathbb{R} \cong \mathbb{C}^2$ . A maximal quartic order  $Q$  (or the quartic field  $K_4$  in which it lies) is *nowhere overramified* if  $Q$  is not overramified at any prime of  $\mathbb{Z}$  (including the archimedean prime).

The significance of being nowhere overramified is as follows. Given an  $S_4$ -quartic field  $K_4$ , let  $K_{24}$  denote its Galois closure. Let  $K_3$  denote a cubic field contained in  $K_{24}$  (the “cubic resolvent field of  $K_4$ ”), and let  $K_6$  be the unique quadratic extension of  $K_3$  such that the Galois closure of  $K_6$  over  $\mathbb{Q}$  is  $K_{24}$ . Then it is a theorem of Heilbronn [31] that the quadratic extension  $K_6/K_3$  is unramified precisely when the quartic field  $K_4$  is nowhere overramified. Conversely, if  $K_3$  is a noncyclic cubic field, and  $K_6$  is an unramified quadratic extension of  $K_3$ , then the Galois closure of  $K_6$  is an  $S_4$ -extension  $K_{24}$  that contains a unique (up to conjugacy) nowhere overramified quartic extension  $K_4$ .

Thus isomorphism classes of nowhere overramified quartic fields  $K_4$  are in canonical bijection with isomorphism classes of pairs  $(K_6, K_3)$  up to isomorphism, where  $K_3$  is a cubic field and  $K_6$  is an unramified quadratic extension. By class field theory, we then see that isomorphism classes of nowhere overramified quartic fields  $K_4$  are in canonical correspondence with pairs  $(K_3, H)$ , where  $K_3$  is a cubic field and  $H$  is an index 2 subgroup of the class group  $\text{Cl}(K_3)$  of  $K_3$ . Of course, index 2 subgroups  $H$  of  $\text{Cl}(K_3)$  are in canonical correspondence with nontrivial characters  $\xi : \text{Cl}(K_3) \rightarrow \{\pm 1\}$  having kernel  $H$ , and so they are thus in correspondence with nontrivial elements of the abelian group dual to  $\text{Cl}_2(K_3)$ , the 2-torsion subgroup of the class group  $\text{Cl}(K_3)$  of  $K_3$ .

We conclude that isomorphism classes of nowhere overramified quartic fields  $K_4$  are in bijective correspondence with pairs  $(K_3, g)$ , where  $K_3$  is a cubic field and  $g$  is a nontrivial element of  $\text{Cl}_2(K_3)^*$ . An analogous argument shows that isomorphism classes of quartic fields  $K_4$  that are not overramified at any *finite* prime are in bijective correspondence with pairs  $(K_3, g)$ , where  $K_3$  is a cubic field and  $g$  is a nontrivial element of  $\text{Cl}_2^+(K_3)^*$ . Restricting the above correspondences to quartic fields having monogenic resolvent cubic fields, in conjunction with Corollary 3.8, now yields parts (a) and (b) of the theorem.

To see the final assertion of the theorem, we recall that if  $C$  is a maximal cubic order, then there is a unique quartic ring  $Q$  (up to isomorphism) that is not an integral domain but has  $C$  as its cubic resolvent ring, namely,  $\mathbb{Z} \oplus C$ . If  $C$  is both maximal and monogenic, then by Corollary 3.8,  $Q$  will correspond to the unique  $\text{GL}_2(\mathbb{Z})$ -equivalence class of binary quartic forms corresponding to the pair  $(\mathbb{Z} \oplus C, C)$ , and any binary quartic form in this equivalence class will have a linear factor over  $\mathbb{Z}$ .  $\square$

Theorem 3.10 implies that if  $I$  and  $J$  are the invariants of a maximal monogenic cubic ring  $C$ , then the set of all  $\mathrm{GL}_2(\mathbb{Z})$ -equivalence classes of integral binary quartic forms having invariants equal to  $(I, J)$  naturally forms a group, namely, the elementary abelian 2-group  $\mathrm{Cl}_2^+(C)^*$ . In particular, the number of  $\mathrm{GL}_2(\mathbb{Z})$ -equivalence classes of integral binary quartic forms having such invariants  $(I, J)$  is always a power of two! We will show in Sections 3.5, 3.6, and 3.7 (in Lemma 3.15, Proposition 3.17, and the proof of Theorem 1.9, respectively) that “most”—indeed a density of  $1/\zeta(2) \approx 60.8\%$  of—pairs  $(I, J)$ , when ordered by height, have this property.

### 3.5 Computations of $p$ -adic densities

A cubic ring  $C$  over  $\mathbb{Z}$  is maximal if and only if for all primes  $p$  it is *maximal at  $p$* , i.e.,  $C \otimes \mathbb{Z}_p$  is the maximal cubic ring over  $\mathbb{Z}_p$  contained in  $C \otimes \mathbb{Q}_p$ . Similarly, a quartic ring  $Q$  is *maximal* (i.e.,  $Q$  is the maximal quartic ring in  $Q \otimes \mathbb{Q}$ ) if and only if for all primes  $p$  it is *maximal at  $p$* , i.e.,  $Q \otimes \mathbb{Z}_p$  is the maximal quartic ring over  $\mathbb{Z}_p$  contained in  $Q \otimes \mathbb{Q}_p$ . We say that a pair  $(Q, C)$ , where  $Q$  is a quartic ring over  $\mathbb{Z}$  and  $C$  is a cubic resolvent ring of  $Q$ , is *strongly maximal* (resp. *strongly maximal at  $p$* ) if both  $Q$  and  $C$  are maximal (resp. maximal at  $p$ ).

In order to determine the average number of 2-torsion elements in the class groups of monogenized cubic fields, when these fields are ordered by their heights, we wish to count: 1) the number of  $F_{\mathbb{Z},1}$ -orbits on the space  $U_{\mathbb{Z},1}$  of integral monic binary cubic forms that correspond to maximal cubic orders  $C$  having bounded height under the bijection of Proposition 3.1; and 2) the number of  $\mathrm{GL}_2(\mathbb{Z})$ -orbits on the space  $V_{\mathbb{Z}}$  of integral binary quartic forms that correspond to strongly maximal pairs  $(Q, C)$  of bounded height under the bijection of Corollary 3.8. To do this, we need to understand, and determine the density of, the set of congruence conditions on  $U_{\mathbb{Z},1}$  and  $V_{\mathbb{Z}}$  that define maximality on the corresponding cubic and quartic rings.

#### Monic binary cubic forms

We first consider the sets  $U_{\mathbb{Z},1}$ ,  $U_{\mathbb{Z}_p,1}$ , and  $U_{\mathbb{F}_p,1}$  of monic binary cubic forms over the integers  $\mathbb{Z}$ , the  $p$ -adic ring  $\mathbb{Z}_p$ , and the residue field  $\mathbb{F}_p$ , respectively. Any form  $g \in U_{\mathbb{Z},1}$  (resp.  $U_{\mathbb{Z}_p,1}$ ,  $U_{\mathbb{F}_p,1}$ ) determines exactly three points in  $\overline{\mathbb{F}}_p$ , obtained by taking the roots in  $\overline{\mathbb{F}}_p$  of the reduction of  $g$  modulo  $p$ . We define the symbol  $(g, p)$  by setting

$$(g, p) = (f_1^{e_1} f_2^{e_2} \cdots),$$

where the  $f_i$ 's indicate the degrees of the fields of definition over  $\mathbb{F}_p$  of the roots of  $g$ , and the  $e_i$ 's indicate the respective multiplicities of these roots. There are thus five possible values for the symbol  $(g, p)$ , namely, (111), (12), (3), (1<sup>2</sup>1), and (1<sup>3</sup>). Furthermore, it is clear that if two monic binary cubic forms  $g_1, g_2$  over  $\mathbb{Z}$  (resp.  $\mathbb{Z}_p, \mathbb{F}_p$ ) are equivalent under a transformation in  $F_{\mathbb{Z},1}$  (resp.  $F_{\mathbb{Z}_p,1}, F_{\mathbb{F}_p,1}$ ), then  $(g_1, p) = (g_2, p)$ .

If  $C(g)$  is the ring corresponding to an integral monic binary cubic form  $g$ , then the symbol  $(g, p)$  conveys precisely the splitting behavior of  $C(g)$  at  $p$ . For example, if  $(g, p) = (111)$  for  $g \in U_{\mathbb{Z},1}$ , then this means that  $p$  splits completely in the cubic ring  $C(g)$ .

Let  $T_p(U_1, (111))$ ,  $T_p(U_1, (12))$ , etc., denote the set of  $g \in U_{\mathbb{Z},1}$  such that  $(g, p) = (111)$ ,  $(g, p) = (12)$ , etc. Let  $\mu = \mu_p$  be the Haar measure on  $U_{\mathbb{Z}_p,1}$  normalized to satisfy  $\mu(U_{\mathbb{Z}_p,1}) = 1$ . For any set  $S$  in  $U_{\mathbb{Z},1}$  that is definable by congruence conditions, let  $\mu(S)$  denote the  $p$ -adic density of the  $p$ -adic closure of  $S$  in  $U_{\mathbb{Z}_p,1}$ . We then determine the densities of the sets  $T_p(U_1, \cdot)$  in the following lemma.

**Lemma 3.11** *We have*

$$\begin{aligned} \mu(T_p(U_1, (111))) &= \frac{1}{6} (p-1)(p-2)/p^2, \\ \mu(T_p(U_1, (12))) &= \frac{1}{2} (p-1)/p, \\ \mu(T_p(U_1, (3))) &= \frac{1}{3} (p^2-1)/p^2, \\ \mu(T_p(U_1, (1^2 1))) &= (p-1)/p^2, \\ \mu(T_p(U_1, (1^3))) &= 1/p^2. \end{aligned}$$

**Proof:** Since the criteria for membership of  $g$  in  $T_p(U_1, \cdot)$  depends only on the residue class of  $g$  modulo  $p$ , it suffices to consider the situation over  $\mathbb{F}_p$ . We examine first  $\mu(T_p(U_1, (111)))$ . The number of unordered

distinct triples in  $\mathbb{F}_p$  is  $\frac{1}{6}p(p-1)(p-2)$ . Furthermore, given such a triple, there is a unique element in  $U_{\mathbb{F}_p,1}$  having this triple as its set of roots. As the total number of forms in  $U_{\mathbb{F}_p,1}$  is equal to  $p^3$ , it follows that  $\mu(T_p(U_1, (111))) = \frac{1}{6}(p-1)(p-2)/p^2$  as given by the lemma.

Similarly, the number of unordered triples of distinct points, one member of which is in  $\mathbb{F}_p$  while the other two are  $\mathbb{F}_p$ -conjugate in  $\mathbb{F}_{p^2}$ , is given by  $\frac{1}{2}p(p^2-p)$ . Also, the number of unordered  $\mathbb{F}_p$ -conjugate triples of distinct points in  $\mathbb{F}_{p^3}$  is  $(p^3-p)/3$ . We thus have  $\mu(T_p(U_1, (12))) = \frac{1}{2}(p-1)/p$  and  $\mu(T_p(U_1, (3))) = \frac{1}{3}(p^2-1)/p^2$ .

Finally, the number of forms in  $U_{\mathbb{F}_p,1}$  having a double root at a point of  $\mathbb{F}_p$  and a single root at a different point of  $\mathbb{F}_p$  is  $p(p-1)$ , and the number of forms in  $U_{\mathbb{F}_p,1}$  having a triple root in  $\mathbb{F}_p$  is  $p$ . Therefore  $\mu(T_p(U_1, (1^21))) = (p-1)/p^2$  and  $\mu(T_p(U_1, (1^3))) = 1/p^2$ .  $\square$

We now state a lemma proved in [9, Section 3].

**Lemma 3.12** *The cubic ring  $C(g)$  corresponding to an integral binary cubic form  $g$  is nonmaximal at  $p$  if and only if any one of the two following conditions are satisfied: (1)  $g$  is a multiple of  $p$ ; or, (2) there exists an integer  $r$  such that  $r \pmod{p}$  is a multiple root of  $g \pmod{p}$ , and furthermore  $g(r)$  is a multiple of  $p^2$ .*

Let  $\mathcal{U}_p(U_1)$  denote the subset of elements  $g \in U_{\mathbb{F}_p,1}$  that are *maximal at  $p$* , i.e., whose associated cubic rings  $C(g)$  are maximal at  $p$ . Similarly, let  $\mathcal{U}_p(U_1, \cdot)$  denote the subset of elements  $g \in T_p(U_1, \cdot)$  such that  $C(g)$  is maximal at  $p$ . In the following lemma, we determine the densities of the sets  $\mathcal{U}_p(U_1, \cdot)$  in  $U_{\mathbb{F}_p,1}$ .

**Lemma 3.13** *We have*

$$\begin{aligned}\mu(\mathcal{U}_p(U_1, (111))) &= \frac{1}{6}(p-1)(p-2)/p^2, \\ \mu(\mathcal{U}_p(U_1, (12))) &= \frac{1}{2}(p-1)/p, \\ \mu(\mathcal{U}_p(U_1, (3))) &= \frac{1}{3}(p^2-1)/p^2, \\ \mu(\mathcal{U}_p(U_1, (1^21))) &= (p-1)^2/p^3, \\ \mu(\mathcal{U}_p(U_1, (1^3))) &= (p-1)/p^3.\end{aligned}$$

**Proof:** If  $g$  is an element of  $T_p(U_1, (111))$ ,  $T_p(U_1, (12))$ , or  $T_p(U_1, (3))$ , then the cubic ring corresponding to  $g$  is clearly maximal at  $p$ , as its discriminant is coprime to  $p$ . Thus  $\mathcal{U}_p(U_1, (111)) = T_p(U_1, (111))$ ,  $\mathcal{U}_p(U_1, (12)) = T_p(U_1, (12))$ , and  $\mathcal{U}_p(U_1, (3)) = T_p(U_1, (3))$ .

Now, if  $g$  is in  $T_p(U_1, (1^21))$  or  $T_p(U_1, (1^3))$ , then it can be brought into the form  $g(x, y) = x^3 + rx^2y + sxy^2 + ty^3$  with  $s \equiv t \equiv 0 \pmod{p}$ , namely, by sending the unique multiple root of  $g \pmod{p}$  to 0  $\pmod{p}$ , via a transformation in  $F_{\mathbb{Z},1}$ . Of all  $g$  in  $T_p(U_1, (1^21))$  or  $T_p(U_1, (1^3))$  that have been rendered in such a form, a proportion of  $1/p$  actually satisfy the congruence  $t \equiv 0 \pmod{p^2}$ . Thus, by Lemma 3.12, a proportion of  $(p-1)/p$  of forms in  $T_p(U_1, (1^21))$  and in  $T_p(U_1, (1^3))$  correspond to cubic rings maximal at  $p$ . The lemma now follows from Lemma 3.11.  $\square$

## Binary quartic forms

We next consider the sets  $V_{\mathbb{Z}}$ ,  $V_{\mathbb{Z}_p}$ , and  $V_{\mathbb{F}_p}$  of binary quartic forms over the integers  $\mathbb{Z}$ , the  $p$ -adic ring  $\mathbb{Z}_p$ , and the residue field  $\mathbb{F}_p$ , respectively. Each form  $f \in V_{\mathbb{Z}}$  (resp.  $V_{\mathbb{Z}_p}$ ,  $V_{\mathbb{F}_p}$ ) determines exactly four points in  $\mathbb{P}_{\mathbb{F}_p}^1$ , obtained by taking the roots of  $f$  reduced modulo  $p$ . Again, for such a form  $f$ , define the symbol  $(f, p)$  by setting

$$(f, p) = (f_1^{e_1} f_2^{e_2} \cdots),$$

where the  $f_i$ 's indicate the degrees of the fields of definition over  $\mathbb{F}_p$  of the roots of  $f$ , and the  $e_i$ 's indicate the respective multiplicities of these roots. There are then eleven possibilities for the value of the symbol  $(f, p)$ , namely, (1111), (112), (13), (22), (4),  $(1^211)$ ,  $(1^22)$ ,  $(1^21^2)$ ,  $(2^2)$ ,  $(1^31)$ , and  $(1^4)$ . Furthermore, it is clear that if two binary quartic forms  $f_1, f_2$  over  $\mathbb{Z}$  (resp.  $\mathbb{Z}_p$ ,  $\mathbb{F}_p$ ) are equivalent under a transformation in  $\text{GL}_2(\mathbb{Z})$  (resp.  $\text{GL}_2(\mathbb{Z}_p)$ ,  $\text{GL}_2(\mathbb{F}_p)$ ), then  $(f_1, p) = (f_2, p)$ .



If  $Q(f)$  is the quartic ring corresponding to an integral binary quartic form  $f$ , then the symbol  $(f, p)$  conveys precisely the splitting behavior of  $Q(f)$  at  $p$ . For example, if  $(f, p) = (1111)$  for  $f \in V_{\mathbb{Z}}$ , then this means that  $p$  splits completely in the quartic ring  $Q(f)$ .

By  $T_p(V, (1111)), T_p(V, (112))$ , etc., let us denote the set of  $f \in V_{\mathbb{Z}}$  such that  $(f, p) = (1111)$ ,  $(f, p) = (112)$ , etc. Now, for any set  $S$  in  $V_{\mathbb{Z}}$  that is definable by congruence conditions, let us denote by  $\mu(S) = \mu_p(S)$  the  $p$ -adic density of the  $p$ -adic closure of  $S$  in  $V_{\mathbb{Z}_p}$ , where we normalize the additive measure  $\mu$  on  $V_{\mathbb{Z}_p}$  so that  $\mu(V_{\mathbb{Z}_p}) = 1$ . Let  $\mathcal{V}_p(V)$  denote the set of elements  $f \in V_{\mathbb{Z}}$  that are *strongly maximal* at  $p$ , i.e., whose associated cubic ring  $C$  is maximal at  $p$  (the associated quartic ring  $Q$  is then automatically maximal at  $p$ ). We see from [5, Section 3.1] that the resolvent cubic  $C$  of a quartic ring  $Q$  can fail to be maximal at  $p$  in two ways: (1)  $Q$  is nonmaximal at  $p$ , or (2)  $Q$  is maximal but overramified at  $p$ . Therefore  $f \in \mathcal{V}_p(V)$  if and only if the corresponding quartic ring is either nonmaximal at  $p$  or it has splitting type  $(1^2 1^2)$ ,  $(2^2)$ , or  $(1^4)$  at  $p$ .

Let  $\mathcal{V}_p(V, \cdot)$  denote the subset of strongly maximal elements  $f \in T_p(V, \cdot)$ . In the next lemma, we determine the densities of the sets  $\mathcal{V}_p(V, \cdot)$ :

**Lemma 3.14** *We have*

$$\begin{aligned} \mu(\mathcal{V}_p(V, (1111))) &= \frac{1}{24}(p+1)(p-1)^2(p-2)/p^4, \\ \mu(\mathcal{V}_p(V, (112))) &= \frac{1}{4}(p+1)(p-1)^2/p^3, \\ \mu(\mathcal{V}_p(V, (13))) &= \frac{1}{3}(p^2-1)^2/p^4, \\ \mu(\mathcal{V}_p(V, (22))) &= \frac{1}{8}(p-1)^2(p^2-p-2)/p^4, \\ \mu(\mathcal{V}_p(V, (4))) &= \frac{1}{4}(p+1)(p-1)^2/p^3, \\ \mu(\mathcal{V}_p(V, (1^2 11))) &= \frac{1}{2}(p-1)^3(p+1)/p^5, \\ \mu(\mathcal{V}_p(V, (1^2 2))) &= \frac{1}{2}(p-1)^3(p+1)/p^5, \\ \mu(\mathcal{V}_p(V, (1^3 1))) &= (p-1)^2(p+1)/p^5, \\ \mu(\mathcal{V}_p(V, (1^2 1^2))) &= 0, \\ \mu(\mathcal{V}_p(V, (2^2))) &= 0, \\ \mu(\mathcal{V}_p(V, (1^4))) &= 0. \end{aligned}$$

**Proof:** If  $f$  is an element of  $T_p(V, (1111))$ ,  $T_p(V, (112))$ ,  $T_p(V, (13))$ ,  $T_p(V, (22))$ , or  $T_p(V, (4))$ , then the discriminant of the resolvent binary cubic form of  $f$  (which is also equal to the discriminant of  $f$ ) is coprime to  $p$ . Thus  $f$  is strongly maximal at  $p$ , and so  $\mathcal{V}_p(V, (1111)) = T_p(V, (1111))$ ,  $\mathcal{V}_p(V, (112)) = T_p(V, (112))$ ,  $\mathcal{V}_p(V, (13)) = T_p(V, (13))$ ,  $\mathcal{V}_p(V, (22)) = T_p(V, (22))$ , and  $\mathcal{V}_p(V, (4)) = T_p(V, (4))$ .

As before, membership of  $f$  in  $T_p(V, \cdot)$  depends only on the reduction of  $f$  modulo  $p$ . Therefore it suffices to consider the situation over  $V_{\mathbb{F}_p}$ . To compute  $\mu(T_p(V, (1111)))$ , note that there are  $\frac{1}{24}(p+1)p(p-1)(p-2)$  quadruples of unordered distinct points in  $\mathbb{P}^1(\mathbb{F}_p)$ . Four distinct points in  $\mathbb{P}^1(\mathbb{F}_p)$  determine the reduction of  $f \in T_p(V, (1111))$  modulo  $p$ , up to scaling by elements in  $\mathbb{F}_p^\times$ , by specifying its roots in  $\mathbb{P}^1(\mathbb{F}_p)$ . As there are  $p^5$  forms in  $V_{\mathbb{F}_p}$ , we obtain  $\mu(T_p(V, (1111))) = \frac{1}{24}(p+1)(p-1)^2(p-2)/p^4$ . The arguments for computing the  $p$ -adic densities of  $T_p(V, (112))$ ,  $T_p(V, (13))$ ,  $T_p(V, (22))$ , and  $T_p(V, (4))$  are similar.

If  $f$  is in  $T_p(V, (1^2 1^2))$ ,  $T_p(V, (2^2))$ , or  $T_p(V, (1^4))$ , then  $f$  is overramified and is not strongly maximal. So  $\mu(\mathcal{V}_p(V, (1^2 1^2)))$ ,  $\mu(\mathcal{V}_p(V, (2^2)))$ , and  $\mu(\mathcal{V}_p(V, (1^4)))$  are each equal to 0.

Suppose that  $f(x, y) = ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4$  belongs to the closure in  $V_{\mathbb{Z}_p}$  of  $T_p(V, (1^2 11))$ ,  $T_p(V, (1^2 2))$ , or  $T_p(V, (1^3 1))$ . Then the reduction of  $f$  modulo  $p$  has a unique multiple root in  $\mathbb{P}^1(\mathbb{F}_p)$ . By a suitable  $\text{GL}_2(\mathbb{Z}_p)$  transformation, we may assume that this multiple root is equal to  $[0, 1]$ , i.e.,  $d \equiv e \equiv 0 \pmod{p}$ . Since  $f$  is not overramified, the reduction modulo  $p$  of the binary quadratic form  $ax^2 + bxy + cy^2$  does not have a double root, i.e.,  $b^2 - 4ac \not\equiv 0 \pmod{p}$ . The reduction modulo  $p^2$  of the resolvent binary cubic form of  $f$  is equal to  $x^3 + cx^2y + (bd - 4ae)xy^2 + e(b^2 - 4ac)y^3$ , and therefore its reduction modulo  $p$  has a double root at  $[0 : 1] \in \mathbb{P}_{\mathbb{F}_p}^1$ . Therefore  $f$  is strongly maximal if and only if  $e(b^2 - 4ac) \not\equiv 0 \pmod{p^2}$ , i.e.,  $e \not\equiv 0 \pmod{p^2}$ . This happens with probability  $1 - 1/p$ . Therefore  $\mu(\mathcal{V}_p(V, (1^3 1))) = (1 - 1/p)\mu(T_p(V, (1^3 1)))$ ,  $\mu(\mathcal{V}_p(V, (1^2 11))) = (1 - 1/p)\mu(T_p(V, (1^2 11)))$ , and  $\mu(\mathcal{V}_p(V, (1^2 2))) = (1 - 1/p)\mu(T_p(V, (1^2 2)))$ . A simple

computation, similar to the computation of  $\mu(T_p(V, (1111)))$ , then shows that these densities are as stated in the lemma.  $\square$

### Summary of $p$ -adic densities for maximality in $U_{\mathbb{Z},1}$ and $V_{\mathbb{Z}}$

Using the fact that  $\mathcal{U}_p(U_1)$  is the disjoint union of the  $\mathcal{U}_p(U_1, \cdot)$ 's and that  $\mathcal{V}_p(V)$  is the disjoint union of the  $\mathcal{V}_p(V, \cdot)$ 's, we arrive at the following lemma.

**Lemma 3.15** *We have*

$$\begin{aligned}\mu(\mathcal{U}_p(U_1)) &= (p^2 - 1)/p^2, \\ \mu(\mathcal{V}_p(V)) &= (p^2 - 1)^2/p^4.\end{aligned}$$

### 3.6 Congruence conditions and uniformity estimates

We now wish to count the number of  $F_{\mathbb{Z},1}$ -equivalence classes of monic integral binary cubic forms of bounded height whose coefficients satisfy a specified finite set of congruence conditions.

Similar to Theorem 2.11, we have:

**Theorem 3.16** *Suppose  $S$  is an  $F_{\mathbb{Z},1}$ -invariant subset of  $U_{\mathbb{Z},1}^{(i)}$  defined by finitely many congruence conditions. Then we have*

$$N(S; X) = N(U_{\mathbb{Z},1}^{(i)}; X) \prod_p \mu_p(S) + O(X^{1/2+\epsilon}). \quad (32)$$

However, the set of forms in  $U_{\mathbb{Z},1}$  corresponding to maximal cubic rings is defined by infinitely many congruence conditions, namely, by congruence conditions modulo  $p^2$  for every prime  $p$ . Similarly, the set of forms in  $V_{\mathbb{Z}}$  corresponding to strongly maximal pairs of rings is defined by infinitely many congruence conditions, again by congruence conditions modulo  $p^2$  for every prime  $p$ . To prove that versions of (24) and (32) still hold for such sets, we require a uniform estimate on the error terms in (24) and (32) as the number of congruence conditions on  $S$  increases.

Let  $\mathcal{W}_p(U_1)$  denote the set of elements in  $U_{\mathbb{Z},1}$  that are not maximal at  $p$ . Let  $\mathcal{W}_p(V)$  denote the set of elements in  $V_{\mathbb{Z}}$  that are not strongly maximal at  $p$ . Then we require the following estimates.

**Proposition 3.17**  $N(\mathcal{W}_p(U_1); X) = O(X^{5/6}/p^{3/2})$ , where the implied constant is independent of  $p$ .

**Proof:** In [2, Lemma 2] it is shown that the number of isomorphism classes of cubic rings that are not maximal at  $p$  and have discriminant less than  $X$  is bounded by  $O(X/p^2)$ , where the implied constant is independent of  $p$ . Now the height of a monogenized cubic ring is at least half of the absolute value of its discriminant; therefore, using the fact that a cubic ring can have at most 12 monogenizations (see [25] and [27]), we obtain that  $N(\mathcal{W}_p(U_1); X) = O(X/p^2)$ .

Next recall that the set  $\mathcal{S} = \{X^3 + rX^2Y + sXY^2 + tY^3 : r \in \{-1, 0, 1\}; s, t \in \mathbb{Z}\}$ , defined in (27), is a fundamental domain for the action of  $F_{\mathbb{Z},1}$  on  $U_{\mathbb{Z},1}$ . Hence  $N(\mathcal{W}_p(U_1); X)$  is equal to the number of forms in  $\mathcal{S}$  having height less than  $X$  that correspond to cubic rings that are not maximal at  $p$ . Now Lemma 3.12 implies that for fixed integer values of  $r$  and  $s$ , there exist at most two values of  $t \pmod{p^2}$  such that the cubic ring corresponding to the form  $f(X, Y) = X^3 + rX^2Y + sXY^2 + tY^3$  is not maximal at  $p$ . Furthermore, by the proof of Lemma 3.3, if  $f \in \mathcal{S}$  and  $H(f) < X$ , then  $s = O(X^{1/3})$  and  $t = O(X^{1/2})$ .

Therefore, we obtain the estimate

$$N(\mathcal{W}_p(U_1); X) = O(X^{1/3}) \cdot O(\max\{2, X^{1/2}/p^2\}) = O(X^{1/3} + X^{5/6}/p^2),$$

which coupled with the earlier estimate  $N(\mathcal{W}_p(U_1); X) = O(X/p^2)$  yields the result.  $\square$

We similarly have

**Proposition 3.18**  $N(\mathcal{W}_p(V); X) = O(X^{5/6}/p^{5/3})$ , where the implied constant is independent of  $p$ .

The proof of Proposition 3.18 is significantly more difficult than that of Proposition 3.17. Its proof will be one of the primary goals of Section 4.

### 3.7 Proof of the main theorem (Theorem 1.9)

In this subsection, we compute the mean number of 2-torsion elements in the class groups of monogenized cubic fields when these monogenized cubic fields are ordered by height. We thus prove Theorem 1.9.

Let  $\mathcal{U}(U_1^{(0)})$  and  $\mathcal{U}(U_1^{(1)})$  denote the subset of elements in  $U_{\mathbb{Z},1}$  corresponding to maximal cubic rings that have positive and negative discriminant, respectively.

**Proposition 3.19** *We have*

$$\begin{aligned} \text{a) } \lim_{X \rightarrow \infty} \frac{N(\mathcal{U}(U_1^{(0)}); X)}{X^{5/6}} &= \frac{8}{135\zeta(2)}; \\ \text{b) } \lim_{X \rightarrow \infty} \frac{N(\mathcal{U}(U_1^{(1)}); X)}{X^{5/6}} &= \frac{32}{135\zeta(2)}. \end{aligned}$$

**Proof:** We only prove part (a) of the proposition as the proof of part (b) is identical. Let  $\mathcal{U}_p(U_1^{(0)})$  be the set of elements in  $U_{\mathbb{Z},1}^{(0)}$  that are maximal at  $p$ . Then by Proposition 3.4, Lemma 3.15 and Theorem 3.16, it follows that for any fixed positive integer  $Y$ , we have

$$\lim_{X \rightarrow \infty} \frac{N(\cap_{p < Y} \mathcal{U}_p(U_1^{(0)}); X)}{X^{5/6}} = \frac{8}{135} \prod_{p < Y} (1 - p^{-2}).$$

Letting  $Y$  tend to infinity, we obtain that

$$\limsup_{X \rightarrow \infty} \frac{N(\mathcal{U}(U_1^{(0)}); X)}{X^{5/6}} = \frac{8}{135} \prod_p (1 - p^{-2}) = \frac{8}{135\zeta(2)}. \quad (33)$$

To obtain a lower bound for  $N(\mathcal{U}(U_1^{(0)}); X)$ , we note that

$$\bigcap_{p < Y} \mathcal{U}_p(U_1^{(0)}) \subset \left( \mathcal{U}(U_1^{(0)}) \cup \bigcup_{p > Y} \mathcal{W}_p(U_1) \right).$$

Hence, by the uniformity estimate of Proposition 3.17, we obtain

$$\lim_{X \rightarrow \infty} \frac{N(\mathcal{U}(U_1^{(0)}); X)}{X^{5/6}} \geq \frac{8}{135} \prod_{p < Y} (1 - p^{-2}) - O\left(\sum_{p \geq Y} p^{-3/2}\right).$$

Letting  $Y$  tend to infinity completes the proof.  $\square$

Identically, by Proposition 2.1, Lemma 3.15, Theorem 2.11, and Proposition 3.18, if we denote the set of elements in  $V_{\mathbb{Z}}^{(i)}$  corresponding to strongly maximal pairs of rings  $(Q, C)$  by  $\mathcal{V}(V^{(i)})$ , then we have

$$\lim_{X \rightarrow \infty} \frac{N(\mathcal{V}(V^{(i)}); X)}{X^{5/6}} = \frac{4\zeta(2)}{135} \prod_p (1 - p^{-2})^2 = \frac{4}{135\zeta(2)} \quad (34)$$

for  $i = 0, 2+$ , and  $2-$ , and

$$\lim_{X \rightarrow \infty} \frac{N(\mathcal{V}(V^{(1)}); X)}{X^{5/6}} = \frac{32\zeta(2)}{135} \prod_p (1 - p^{-2})^2 = \frac{32}{135\zeta(2)}. \quad (35)$$

By Theorem 3.10, we see that

$$\begin{aligned}
\sum_{\substack{H(K) < X \\ \text{Disc}(K) > 0}} (\#\text{Cl}_2(K) - 1) &= N(\mathcal{V}(V^{(0)}); X), \\
\sum_{\substack{H(K) < X \\ \text{Disc}(K) < 0}} (\#\text{Cl}_2(K) - 1) &= N(\mathcal{V}(V^{(1)}); X), \\
\sum_{\substack{H(K) < X \\ \text{Disc}(K) > 0}} (\#\text{Cl}_2^+(K) - 1) &= N(\mathcal{V}(V^{(0)}); X) + N(\mathcal{V}(V^{(2+)}); X) + N(\mathcal{V}(V^{(2-)}); X),
\end{aligned}$$

where  $K = (K, x)$  ranges over monogenized cubic fields. We therefore have

$$\begin{aligned}
\lim_{X \rightarrow \infty} \frac{\sum_{\substack{H(K) < X \\ \text{Disc}(K) > 0}} \#\text{Cl}_2(K)}{\sum_{\substack{H(K) < X \\ \text{Disc}(K) > 0}} 1} &= 1 + \frac{4/(135\zeta(2))}{8/(135\zeta(2))} = 1.5, \\
\lim_{X \rightarrow \infty} \frac{\sum_{\substack{H(K) < X \\ \text{Disc}(K) < 0}} \#\text{Cl}_2(K)}{\sum_{\substack{H(K) < X \\ \text{Disc}(K) < 0}} 1} &= 1 + \frac{32/(135\zeta(2))}{32/(135\zeta(2))} = 2, \\
\lim_{X \rightarrow \infty} \frac{\sum_{\substack{H(K) < X \\ \text{Disc}(K) > 0}} \#\text{Cl}_2^+(K)}{\sum_{\substack{H(K) < X \\ \text{Disc}(K) > 0}} 1} &= 1 + \frac{12/(135\zeta(2))}{8/(135\zeta(2))} = 2.5.
\end{aligned}$$

This proves the weaker version of Theorem 1.9, where we average over all monogenized cubic fields without any specified local splitting conditions.

To complete the proof of Theorem 1.9 also in the case where finitely many splitting conditions are imposed, we note that if a single splitting type  $\sigma$  is imposed on the cubic fields we are considering at a given prime  $p$ , then the Euler factor at  $p$  in (33) changes from  $\mu_p(\mathcal{U}_p(U_1)) = (1 - p^{-2})$  to  $\mu_p(\mathcal{U}_p(U_1, \sigma))$ ; meanwhile, the Euler factor at  $p$  in (34) changes from  $\mu_p(\mathcal{V}_p(V)) = (1 - p^{-2})^2$  to

$$\sum_{\theta \in R^{-1}(\sigma)} \mu_p(\mathcal{V}_p(V, \theta)), \tag{36}$$

where  $R$  is the natural map from quartic splitting types to cubic splitting types, namely: given a strongly maximal pair of rings  $(Q, C)$ , the splitting type  $\theta$  of  $Q$  at  $p$  determines the splitting type  $R(\theta)$  of  $C$  at  $p$ .

Thus to complete the proof of Theorem 1.9, it suffices to prove that the equality

$$\frac{\mu_p(\mathcal{U}_p(U_1, \sigma))}{\mu_p(\mathcal{U}_p(U_1))} = \frac{\sum_{\theta \in R^{-1}(\sigma)} \mu_p(\mathcal{V}_p(V, \theta))}{\mu_p(\mathcal{V}_p(V))} \tag{37}$$

of  $p$ -adic density ratios holds for all values of  $p$  and  $\sigma$ .

To prove (37), we first determine  $R(\theta)$  in terms of  $\theta$  using the following well-known fact (see, e.g., [33] for a proof).

$\sigma$	$\frac{\mu_p(\mathcal{U}_p(U_1, \sigma))}{\mu_p(\mathcal{U}_p(U_1))}$	$R^{-1}(\sigma)$	$\sum_{\theta \in R(\sigma)} \frac{\mathcal{V}_p(V, \theta)}{\mu_p(\mathcal{V}_p(V))}$
(111)	$\frac{p-2}{6(p+1)}$	$\{(1111), (22)\}$	$\frac{p-2}{24(p+1)} + \frac{p-2}{8(p+1)} = \frac{p-2}{6(p+1)}$
(12)	$\frac{p}{2(p+1)}$	$\{(112), (4)\}$	$\frac{p}{4(p+1)} + \frac{p}{4(p+1)} = \frac{p}{2(p+1)}$
(3)	$\frac{1}{3}$	$\{(13)\}$	$\frac{1}{3}$
(1 <sup>2</sup> 1)	$\frac{p-1}{p(p+1)}$	$\{(1^211), (1^22)\}$	$\frac{p-1}{2p(p+1)} + \frac{p-1}{2p(p+1)} = \frac{p-1}{p(p+1)}$
(1 <sup>3</sup> )	$\frac{1}{p(p+1)}$	$\{(1^31)\}$	$\frac{1}{p(p+1)}$

Table 2: Demonstration of the equality of  $p$ -adic density ratios for  $U_1$  and  $V$

**Lemma 3.20** *For any field  $F$ , there is a natural bijection between isomorphism classes of étale  $F$ -algebras of degree  $n$  and isomorphism classes, under conjugation within  $S_n$ , of continuous homomorphisms  $\psi : \text{Gal}(F) \rightarrow S_n$ , where  $\text{Gal}(F)$  denotes the absolute Galois group of  $F$ .*

Now consider the canonical surjection  $\lambda : S_4 \rightarrow S_3$  whose kernel consists of the identity and the double transpositions in  $S_4$ . If  $L$  is the degree four étale algebra over  $\mathbb{Q}_p$  corresponding to the map  $\psi : \text{Gal}(\mathbb{Q}_p) \rightarrow S_4$ , then one shows that the cubic resolvent  $K$  of  $L$  corresponds to the map  $\lambda \circ \psi : \text{Gal}(\mathbb{Q}_p) \rightarrow S_3$  (see [4]).

Thus, for example, if the splitting type of  $Q$  at  $p$  is (22), then the image of the map  $\psi : \text{Gal}(\mathbb{Q}_p) \rightarrow S_4$  corresponding to  $Q \otimes \mathbb{Q}_p$  is a cyclic group  $C_2$  of order 2, where the nontrivial element of  $C_2$  is a double transposition. Hence the image of the composite map  $\lambda \circ \psi : G_{\mathbb{Q}_p} \rightarrow S_3$  is trivial and the splitting type of  $p$  in  $C$  is (111). Therefore  $R((22)) = (111)$ . The computations of  $R(\theta)$  for other splitting types  $\theta$  are similar.

Now suppose  $p$  is a fixed prime and  $\sigma = (111)$ . Then the set  $R^{-1}(\sigma)$  is equal to  $\{(1111), (22)\}$ . By Lemmas 3.13 and 3.15, we see that

$$\frac{\mu_p(\mathcal{U}_p(U_1, \sigma))}{\mu_p(\mathcal{U}_p(U_1))} = \frac{p-2}{6(p+1)},$$

while from Lemmas 3.14 and 3.15, we see that

$$\frac{\mu_p(\cup_{\theta \in R(\sigma)} \mathcal{V}_p(V, \theta))}{\mu_p(\mathcal{V}_p(V))} = \frac{p-2}{24(p+1)} + \frac{p-2}{8(p+1)} = \frac{p-2}{6(p+1)}.$$

The computations for other values of  $\sigma$  are identical and are summarized in Table 2. This completes the proof of Theorem 1.9.

### 3.8 Proofs of auxiliary lemmas

In this subsection, we complete the proofs of various auxiliary lemmas whose proofs were deferred. We begin by determining the order of the stabilizer in  $\text{GL}_2(\mathbb{R})$  of binary quartic forms in  $V_{\mathbb{R}}$ , thus proving Lemma 2.2.

**Proof of Lemma 2.2:** We start by computing the size of the stabilizer in  $\text{PGL}_2(\mathbb{R})$  of binary quartic forms in  $V_{\mathbb{R}}$  having nonzero discriminant, where  $\text{PGL}_2(\mathbb{R})$  acts on  $V_{\mathbb{R}}$  via the “twisted action” defined in (31).

By restricting the results of §3.3 to the case  $R = \mathbb{R}$ , we obtain an injective map  $\rho : \text{PGL}_2(\mathbb{R}) \rightarrow \text{SL}_3(\mathbb{R})$ , and a map  $\phi : V_{\mathbb{R}} \rightarrow W_{\mathbb{R}}$  which is equivariant with respect to  $\rho$ . Let  $f$  be an element in  $V_{\mathbb{R}}$  having nonzero discriminant and let  $\phi(f) = (A_1, B_f)$ . It is clear that  $\gamma \in \text{PGL}_2(\mathbb{R})$  stabilizes  $f \in V_{\mathbb{R}}$  if and only if  $\rho(\gamma)$  stabilizes  $(A_1, B_f)$ , and  $\rho(\gamma)$  stabilizes  $(A_1, B_f)$  if and only if  $(\text{id}, \rho(\gamma)) \in G_{\mathbb{R}} = \text{GL}_2(\mathbb{R}) \times \text{SL}_3(\mathbb{R})$  stabilizes  $(A_1, B_f)$ .

An element  $(A, B) \in W_{\mathbb{R}}$  that is nondegenerate (i.e., the discriminant of the cubic resolvent form of  $(A, B)$  is nonzero) determines four points  $P_1, P_2, P_3$ , and  $P_4$  in  $\mathbb{P}^2(\mathbb{C})$ ; these are the four intersection points in  $\mathbb{P}^2(\mathbb{C})$  of the two conics defined by  $A$  and  $B$ . It follows from [4] that the stabilizer of  $(A, B)$  in  $G_{\mathbb{R}}$  can be canonically realized as a subgroup of the permutation group  $S_4$  on the set of points  $\{P_1, P_2, P_3, P_4\}$ . Furthermore, an element  $\gamma \in G_{\mathbb{R}}$  stabilizing  $(A_1, B_f)$  has an inverse image in  $\mathrm{PGL}_2(\mathbb{R})$  under  $\rho$  if and only if  $\gamma$  is contained in the kernel of the canonical surjection  $\lambda : S_4 \rightarrow S_3$  so that  $\gamma$  is either the identity or a double transposition.

If the splitting type of  $f$  at  $\infty$  is equal to (112) (resp. (1111), (22)), then the stabilizer of  $(A_1, B_f)$  in  $\mathrm{PGL}_2(\mathbb{R})$  is a subgroup  $C_2 \times C_2$  (resp.  $S_4$ ,  $D_4$ ) of  $S_4$  which contains 1 (resp. 3, 3) double transpositions. Now if  $\gamma \in \mathrm{GL}_2(\mathbb{R})$  stabilizes  $f \in V_{\mathbb{R}}$  under the action defined in Equation (3), then since  $I(\gamma \cdot f) = (\det \gamma)^4 I(f)$  and  $J(\gamma \cdot f) = (\det \gamma)^6 J(f)$ , the determinant  $\det \gamma \in \mathbb{R}$  of  $\gamma$  is equal to  $\pm 1$ . Hence the image of  $\gamma$  in  $\mathrm{PGL}_2(\mathbb{R})$  also stabilizes  $f$ . Since there are two elements in the center of  $\mathrm{GL}_2(\mathbb{R})$  that stabilize  $f$ , the size of the stabilizer in  $\mathrm{GL}_2(\mathbb{R})$  of an element  $v \in V_{\mathbb{R}}^{(i)}$  is 4 for  $i = 1$  and 8 for  $i = 0$  or 2, as desired.  $\square$

Our next goal is to prove Lemma 2.4, which bounds the number of  $\mathrm{GL}_2(\mathbb{Z})$ -equivalence classes of integral binary quartic forms having bounded height and nontrivial stabilizer in  $\mathrm{GL}_2(\mathbb{Z})$ . We first require:

**Lemma 3.21** *The number of  $\mathrm{GL}_2(\mathbb{Z})$ -equivalence classes of binary quartic forms  $f$  such that  $H(f) < X$ , and the binary cubic resolvent form of  $f$  is reducible, is bounded by  $O(X^{3/4+\epsilon})$ .*

**Proof:** Let  $g$  be an element of  $U_{\mathbb{Z},1}$  that has discriminant  $n$  and is reducible over  $\mathbb{Z}$ , and let  $C$  be the corresponding cubic ring. Since  $C$  is a cubic ring lying in a cubic  $\mathbb{Q}$ -algebra of the form  $\mathbb{Q} \oplus F$ , where  $F$  is a quadratic  $\mathbb{Q}$ -algebra, by [5, Proof of Lemma 12] the number of quartic rings, up to isomorphism, having  $C$  as a cubic resolvent ring is bounded by  $O(n^{1/4+\epsilon})$ .

Now the height of a cubic ring having discriminant  $n$  is at least  $n/2$ . Moreover, we see from Lemma 3.3 that the number of monogenized cubic rings having height less than  $X$  that correspond to reducible integral monic binary cubic forms is bounded by  $O(X^{1/2+\epsilon})$ . We conclude from Corollary 3.8 that the number of  $\mathrm{GL}_2(\mathbb{Z})$ -equivalence classes of binary quartic forms  $f$  such that  $H(f) < X$ , and the cubic resolvent form of  $f$  is reducible, is bounded by  $O(X^{1/2+\epsilon} \cdot X^{1/4+\epsilon}) = O(X^{3/4+\epsilon})$ .  $\square$

**Proof of Lemma 2.4:** Suppose an integral binary quartic form  $f$  has a stabilizer of size greater than or equal to 2 in  $\mathrm{PGL}_2(\mathbb{Z})$ , and let  $(Q, C)$  be the corresponding pair of rings. Then  $Q$  has a nontrivial automorphism over  $\mathbb{Z}$  and therefore cannot be an order in an  $S_4$ - or an  $A_4$ -quartic field. If  $Q$  is an order in a  $D_4$ -,  $C_4$ -, or  $V_4$ -quartic field, or is a quartic ring lying in a direct sum of two quadratic  $\mathbb{Q}$ -algebras, then the cubic resolvent ring  $C$  of  $Q$  cannot be an integral domain; the desired bound in this case then follows from Lemma 3.21.

Thus we may assume  $Q$  is a quartic ring in an étale quartic  $\mathbb{Q}$ -algebra of the form  $F = \mathbb{Q} \oplus K$ , where  $K$  is a cubic field, and  $\mathrm{Disc}(Q) = k^2 \mathrm{Disc}(K)$  for some  $k \in \mathbb{Z}_{>0}$ . As  $Q$  has nontrivial automorphisms,  $K$  must be a  $C_3$ -cubic field. Thus  $Q$  has a cyclic group of order 3 as its automorphism group, as does  $C$ . The number of possibilities for  $C$  with discriminant at most  $X$  is then  $O(X^{1/2})$  by [10, Theorem 1]. Once  $C$  has been fixed,  $K$  is determined; furthermore, it follows from a result of Nakagawa [37] that the number of orders of index  $k$  in the ring of integers in the étale  $Q$ -algebra  $F = \mathbb{Q} \oplus K$  is at most  $O(k^{1/2+\epsilon}) = O(X^{1/4+\epsilon})$ , where the implied constant is independent of  $F$ . We conclude that the total number of possibilities for the pair  $(Q, C)$  having discriminant at most  $X$  is  $O(X^{1/2} \cdot X^{1/4+\epsilon}) = O(X^{3/4+\epsilon})$  in this case. The same estimate thus also holds for height as in the proof of Lemma 3.21, yielding the desired conclusion.  $\square$

Finally, we determine when a pair of invariants  $(I, J) \in \mathbb{Z} \times \mathbb{Z}$  is eligible thus proving Lemma 2.9.

**Proof of Lemma 2.9:** If an integral binary quartic form has invariants equal to  $I$  and  $J$ , then its cubic resolvent form also has invariants equal to  $I$  and  $J$ . Therefore, for an integer pair  $(I, J)$  to be eligible, it must satisfy one of the congruences listed in Lemma 3.2. Conversely, suppose an integer pair  $(I, J)$  satisfies one of the congruences of Lemma 3.2. By Lemma 3.2, there exists an integral monic binary cubic form  $g(X, Y) = X^3 + rX^2Y + sXY^2 + tY^4$  having invariants equal to  $I$  and  $J$ . One then checks that the cubic resolvent form of the binary quartic form  $f(x, y) = x^3y + rx^2y^2 + sxy^3 + ty^4$  is equal to  $g$ , and so  $f$  has invariants equal to  $I$  and  $J$ . Therefore the pair  $(I, J)$  is eligible.  $\square$

## 4 The mean number of 2-torsion elements in the class groups of submonogenized cubic fields

In the previous section we determined the mean number of 2-torsion elements in the class groups of monogenized cubic fields when these monogenized cubic fields are ordered by height. We now consider “submonogenized cubic fields”. We define a *submonogenized cubic field of index  $n$*  (or, an  *$n$ -monogenized cubic field*) to be a cubic field  $F$  together with an element  $x$  in the ring of integers  $\mathcal{O}_K$  of  $K$  such that the index of  $\mathbb{Z}[x]$  in  $\mathcal{O}_K$  is equal to  $n$ . If  $g$  denotes the characteristic polynomial of  $\times x : \mathcal{O}_K \rightarrow \mathcal{O}_K$ , then we define the *height* of the  $n$ -monogenized field  $(K, x)$  by

$$H(K, x) := n^2 H(I(g), J(g)). \quad (38)$$

We say that two  $n$ -monogenized cubic rings  $(C, x)$  and  $(C', x')$  are *isomorphic* if  $C$  is isomorphic to  $C'$  and there exists  $u \in \mathbb{Z}$  such that under this isomorphism  $x$  is mapped to  $x' + u$ .

In this section, we determine the mean number of 2-torsion elements in the class groups of submonogenized cubic fields having height bounded by  $X$  and index bounded by  $X^\delta$ , thus proving Theorem 1.10. Surprisingly, these means are different from those obtained in Theorem 1.9 for monogenized cubic fields; as expected, they instead coincide with the averages for the number of 2-torsion elements in the class groups of cubic fields when they are ordered by discriminant (see [5]). Theorems 1.9 and 1.10 demonstrate that, on average, the monogenicity of a ring has a clear altering effect on the class group.

This section is organized as follows. In §4.1, we determine the number of isomorphism classes of submonogenized cubic rings  $C$  having bounded height and index. In §4.2, we similarly determine the number of isomorphism classes of pairs  $(Q, C)$ , where  $Q$  is a quartic ring and  $C$  is a submonogenized cubic resolvent ring of  $Q$  having bounded height and index. The layout of each of these subsections is largely parallel to that of Section 2. In §4.3, §4.4, and §4.5 we then prove, as in Sections §3.5 and §3.6, the various results needed to perform a sieve (analogous to the one in §3.7) to maximal orders. The proof of Theorem 1.10 is then completed in §4.7.

### 4.1 The number of submonogenized cubic rings having bounded height and bounded index

In this subsection, we count the number of (isomorphism classes of) submonogenized cubic rings having height bounded by  $X$  and index bounded by  $X^\delta$ , where  $\delta \leq 1/4$  is a positive constant. More precisely, we prove the following theorem:

**Theorem 4.1** *Let  $N_3^{(0)}(X, \delta)$  (resp.  $N_3^{(1)}(X, \delta)$ ) denote the number of submonogenized rings  $(C, x)$  of index  $n$  with positive (resp. negative) discriminant such that  $H(C, x) < X$  and  $n < X^\delta$ . Then we have*

$$\begin{aligned} \text{(a)} \quad N_3^{(0)}(X, \delta) &= \frac{4}{45} X^{5/6+2\delta/3} + O(X^{5/6}); \\ \text{(b)} \quad N_3^{(1)}(X, \delta) &= \frac{16}{45} X^{5/6+2\delta/3} + O(X^{5/6}). \end{aligned}$$

#### 4.1.1 The parametrization of submonogenized cubic rings of index $n$

For an integer  $n > 0$ , let  $U_{\mathbb{Z}, n}$  denote the space of all integral binary cubic forms whose  $x^3$  coefficient is equal to  $n$ . To an element  $g(x, y) = nx^3 + bx^2y + cxy^2 + dy^3 \in U_{\mathbb{Z}, n}$ , we associate the  $n$ -monogenized cubic ring  $(\mathbb{Z}[X, Y]/(X^3 + bX^2 + ncX + n^2d, X - nY), X)$ . As the  $n$ -monogenized cubic rings  $(C, X)$  and  $(C, X + u)$  are isomorphic, it is also natural to consider the corresponding integral binary cubic forms  $g(x, y)$  and  $g(x + u, y)$  in  $U_{\mathbb{Z}, n}$  to be isomorphic for all  $u \in \mathbb{Z}$ . We thus consider the action of the group  $F_{\mathbb{Z}, 1}$  on  $U_{\mathbb{Z}, n}$ , defined as before by  $\gamma \cdot g(x, y) = g((x, y) \cdot \gamma)$ . We then have the following proposition generalizing Proposition 3.1:

**Proposition 4.2** *The  $F_{\mathbb{Z}, 1}$ -orbits on  $U_{\mathbb{Z}, n}$  are in bijective correspondence with isomorphism classes of  $n$ -monogenized cubic rings over  $\mathbb{Z}$ .*

Let  $U_{\mathbb{R}}$  denote the set of all binary cubic forms with coefficients in  $\mathbb{R}$ . Let  $U_{\mathbb{R},+} \subset U_{\mathbb{R}}$  consist of those elements  $g(x, y) = ax^3 + bx^2y + cxy^2 + dy^3$  in  $U_{\mathbb{R}}$  such that  $a > 0$ . As before, the group

$$F_{\mathbb{R}} = \left\{ \begin{pmatrix} p & \\ q & r \end{pmatrix} : p, q, r \in \mathbb{R}, p, r > 0 \right\}$$

acts naturally on  $U_{\mathbb{R}}$  (and  $U_{\mathbb{R},+}$ ) under this action.

Given a binary cubic form  $g(x, y) = ax^3 + bx^2y + cxy^2 + dy^3$  in  $U_{\mathbb{R},+}$ , it is  $F_{\mathbb{R}}$ -equivalent to the monic binary cubic form  $g'(x, y) = x^3 + \frac{b}{a^{1/3}}x^2y + a^{1/3}cxy^2 + ady^3$ . Furthermore, if  $\gamma \in F_{\mathbb{R},1}$ , then the  $x^3$  coefficient of  $\gamma \cdot g$  remains equal to  $a$ . It is thus natural to define the quantities

$$\begin{aligned} I(g) &:= \frac{b^2}{a^{2/3}} - 3a^{1/3}c, \\ J(g) &:= -2b^3/a + 9bc - 27ad, \\ a(g) &:= a, \end{aligned} \tag{39}$$

which are then invariant under the action of  $F_{\mathbb{R},1} \subset F_{\mathbb{R}}$  on  $U_{\mathbb{R}}$  (compare with (26) in §3.1). As usual, the discriminant of  $g$  can be written in terms of  $I(g)$  and  $J(g)$  via the same formula, namely,  $\Delta(g) = (4I(g)^3 - J(g)^2)/27$ . We again define the *height*  $H(g)$  of  $g \in U_{\mathbb{R},+}$  by

$$H(g) := H(I, J) = \max\{|I(g)|^3, J(g)^2/4\}. \tag{40}$$

If  $(C, x)$  is an  $n$ -monogenized cubic ring and  $g$  is a corresponding integral binary cubic form, then we define the invariants  $I$ ,  $J$ , and  $a$  of  $(C, x)$  by  $I(C, x) := I(g)$ ,  $J(C, x) := J(g)$ , and  $a(C, x) := a(g) = n$ . In particular, comparing with (38), we see that  $H(C, x) = H(g)$ .

Let  $U_{\mathbb{R},+}^{(0)}$  (resp.  $U_{\mathbb{R},+}^{(1)}$ ) denote the subset of elements in  $U_{\mathbb{R},+}$  having positive (resp. negative) discriminant, and let  $U_{\mathbb{Z},+}^{(0)}$  (resp.  $U_{\mathbb{Z},+}^{(1)}$ ) denote the subset of integral elements. In the remainder of this subsection, we prove the following result, which combined with Proposition 4.2 immediately yields Theorem 4.1.

**Proposition 4.3** *For an  $F_{\mathbb{Z},1}$ -invariant set  $S$ , let  $N(S; X, \delta)$  denote the number of  $F_{\mathbb{Z},1}$ -equivalence classes of irreducible elements  $g \in S$  satisfying  $H(g) < X$  and  $0 < a(g) < X^\delta$ . Then, for any positive constant  $\delta \leq 1/4$ , we have*

$$\begin{aligned} \text{(a)} \quad N(U_{\mathbb{Z}}^{(0)}; X, \delta) &= \frac{4}{45} \cdot X^{5/6+2\delta/3} + O(X^{5/6}); \\ \text{(b)} \quad N(U_{\mathbb{Z}}^{(1)}; X, \delta) &= \frac{16}{45} \cdot X^{5/6+2\delta/3} + O(X^{5/6}). \end{aligned}$$

#### 4.1.2 Reduction theory

We first develop the reduction theory necessary to construct a fundamental domain for the action of  $F_{\mathbb{Z},1}$  on  $U_{\mathbb{R},+}$ . Our methods are quite similar to those in §2.1; we start by describing a fundamental set for the action of  $F_{\mathbb{R}}$  on  $U_{\mathbb{R},+}$ .

Let  $I_0, J_0 \in \mathbb{R} \times \mathbb{R}$  be such that  $4I_0^3 - J_0^2 \neq 0$ . Then there exists exactly one  $F_{\mathbb{R}}$ -equivalence class in  $U_{\mathbb{R},+}$  with invariants  $I_0$  and  $J_0$ . The analogous result for  $\text{SL}_2(\mathbb{C})$ -equivalence classes in  $V_{\mathbb{C}}$  was proven in the proof of Proposition 2.8, and a very similar argument applies to the case of  $F_{\mathbb{R}}$ -equivalence classes in  $U_{\mathbb{R},+}$ .

As with the action of  $\text{SL}_2(\mathbb{C})$  on  $V_{\mathbb{C}}$ , a fundamental set  $L_U^{(0)}$  (resp.  $L_U^{(1)}$ ) for the action of  $F_{\mathbb{R}}$  on  $U_{\mathbb{R},+}^{(0)}$  (resp.  $U_{\mathbb{R},+}^{(1)}$ ) may be constructed by choosing one form  $g$  in  $U_{\mathbb{R},+}^{(0)}$  (resp.  $U_{\mathbb{R},+}^{(1)}$ ) for each  $(I, J)$  such that  $H(I, J) = 1$  and  $4I^3 - J^2 > 0$  (resp.  $4I^3 - J^2 < 0$ ). Explicit constructions of such sets  $L_U^{(i)}$  are provided in Table 3. As in §2.1, the key fact about these  $L_U^{(i)}$  that we require is that the absolute values of all the coefficients of all the forms in our  $L_U^{(i)}$  are uniformly bounded.

We also need the following lemma whose proof is elementary:



$$\begin{aligned}
L_U^{(0)} &= \left\{ x^3 - \frac{1}{3}xy^2 - \frac{t}{27}y^3 : -2 \leq t \leq 2 \right\} \\
L_U^{(1)} &= \left\{ x^3 - \frac{s}{3}xy^2 - \frac{1}{27}y^3 : -1 \leq s \leq 1 \right\} \cup \left\{ x^3 + \frac{1}{3}xy^2 - \frac{t}{27}y^3 : -2 \leq t \leq 2 \right\}
\end{aligned}$$

Table 3: Explicit constructions of fundamental sets  $L_U^{(i)}$  for  $\mathbb{F}_{\mathbb{R}} \backslash U_{\mathbb{R}}^{(i)}$

**Lemma 4.4** *For any  $g \in U_{\mathbb{R},+}$ , the stabilizer of  $g$  in  $F_{\mathbb{R}}$  is trivial.*

Let  $\mathcal{F}_1$  denote the fundamental domain for  $F_{\mathbb{Z},1} \backslash F_{\mathbb{R}}$  given by the elements  $n(u)\alpha(t)\lambda$  with  $n \in N'$ ,  $\alpha \in A_+$ , and  $\lambda \in \Lambda$ , where

$$N' = \left\{ \begin{pmatrix} 1 & \\ u & 1 \end{pmatrix} : -1/2 < u \leq 1/2 \right\}, \quad A_+ = \left\{ \begin{pmatrix} t^{-1} & \\ & t \end{pmatrix} : t > 0 \right\}, \quad \Lambda = \left\{ \begin{pmatrix} \lambda & \\ & \lambda \end{pmatrix} : \lambda > 0 \right\}. \quad (41)$$

Then  $\mathcal{F}_1 L_U^{(i)}$  is a fundamental domain for the action of  $F_{\mathbb{Z},1}$  on  $U_{\mathbb{R},+}^{(i)}$ .

#### 4.1.3 Estimates on reducibility

Let  $\mathcal{R}_X(L_U^{(i)}; \delta)$  consist of the elements  $g$  in the fundamental domain  $\mathcal{F}_1 L_U^{(i)}$  satisfying  $H(g) < X$  and  $a(g) < X^\delta$ . That is, define

$$\mathcal{R}_X(L_U^{(i)}; \delta) := \mathcal{F}_1^{(\delta)} L_U^{(i)} \cap \{w : |H(w)| < X\}$$

where  $\mathcal{F}_1^{(\delta)} \subset \mathcal{F}_1$  consists of those elements  $n\alpha\lambda$  in  $\mathcal{F}_1$  satisfying  $\lambda/X^{\frac{\delta}{3}} \leq t \leq \lambda$ . We prove the following lemma.

**Lemma 4.5** *Let  $\delta \leq 1/4$  be a positive constant. Then the number of integral binary cubic forms in  $\mathcal{R}_X(L_U^{(i)}; \delta)$  that are reducible over  $\mathbb{Z}$  is  $O(X^{3/4+\epsilon})$ .*

**Proof:** If  $g = n\alpha(t)\lambda \in \mathcal{F}_1^{(\delta)}$  and  $v \in L_U^{(i)}$  such that  $gv \in \mathcal{R}_X(L_U^{(i)}; \delta)$ , then this implies  $\lambda = O(X^{1/12})$ . Since  $\delta \leq 1/4$ , we have the estimates  $abc = O(\lambda^9/t^3) = O(X^{3/4})$ ,  $abd = O(\lambda^9/t) = O(X^{3/4})$ ,  $ab = O(\lambda^6/t^4) = O(X^{1/2})$ ,  $ac = O(\lambda^6/t^2) = O(X^{1/2})$ , and  $ad = O(\lambda^6) = O(X^{1/2})$ .

If  $ax^3 + bx^2y + cxy^2 + dy^3$  is a reducible binary cubic form, then it has a linear factor. The estimates on  $abc$ ,  $ab$ , and  $ac$  show that the number of forms in  $\mathcal{R}_X(L_U^{(i)}; \delta)$  with  $d = 0$  is at most  $O(X^{3/4})$ . If  $d \neq 0$ , then the estimates on  $abd$  and  $ad$  show that there are at most  $O(X^{3/4})$  possibilities for the triple  $(a, b, d)$  in  $\mathcal{R}_X(L_U^{(i)}; \delta)$ . Once the coefficients  $a, b, d$  have been fixed, there are then only  $O(X^\epsilon)$  possibilities for an integral linear factor  $px + qy$  of the binary cubic form, because  $p$  must divide  $a$  and  $q$  must divide  $d$ . Since  $a, b, d, p$ , and  $q$  then determine  $c$ , we conclude that the number of reducible integral binary cubic forms in  $\mathcal{R}_X(L_U^{(i)}; \delta)$  is at most  $O(X^{3/4+\epsilon})$ , as desired.  $\square$

#### 4.1.4 Averaging and cutting off the cusp

Let  $F_0$  be a compact set in  $F_{\mathbb{R}}$  that is the closure of some open nonempty set, and assume further that every element in  $F_0$  has determinant greater than 1. Let  $S \subset U_{\mathbb{Z},+}^{(i)}$  be an  $F_{\mathbb{Z},1}$ -invariant set and let  $S^{\text{irr}}$  denote the subset of irreducible forms inside  $S$ . Then we have

$$N(S; X, \delta) = \frac{\int_{h \in F_0} \#\{x \in S^{\text{irr}} \cap \mathcal{F}hL_U^{(i)} : H(x) < X, \ 0 < a(x) < X^\delta\} dh}{\int_{h \in F_0} dh}.$$

Just as in §2.3, we see that if  $C_{F_0} = \int_{h \in F_0} dh$ , then

$$N(S; X, \delta) = \frac{1}{C_{F_0}} \int_{g \in \mathcal{F}_1} \# \{x \in S^{\text{irr}} \cap gF_0L_U^{(i)} : H(x) < X, 0 < a(x) < X^\delta\} dg \quad (42)$$

$$= \frac{1}{C_{F_0}} \int_{g \in N'A'\Lambda} \# \left\{ x \in S^{\text{irr}} \cap n \begin{pmatrix} t^{-1} & \\ & t \end{pmatrix} \lambda F_0L_U^{(i)} : H(x) < X, 0 < a(x) < X^\delta \right\} t^{-2} dn d^\times t d^\times \lambda. \quad (43)$$

Define  $B(n, t, \lambda, X, \delta) := n \begin{pmatrix} t^{-1} & \\ & t \end{pmatrix} \lambda F_0L_U^{(i)} \cap \{x \in U_{\mathbb{R},+}^{(i)} : H(x) < X, 0 < a(x) < X^\delta\}$ . Then

$$N(S^{\text{irr}}; X, \delta) = \frac{1}{C_{F_0}} \int_{g \in N'A'\Lambda} \# \{x \in S^{\text{irr}} \cap B(n, t, \lambda, X, \delta)\} t^{-2} dn d^\times t d^\times \lambda. \quad (44)$$

By our construction of the  $L_U^{(i)}$  and  $F_0$ , all the absolute values of the coefficients of the forms in  $F_0L_U^{(i)}$  are uniformly bounded. Let  $C_1$  and  $C_2$  be constants that bound the absolute values of the  $x^3$  coefficients of all the forms in  $F_0L_U^{(i)}$  from above and below, respectively. Let  $C_3$  be a constant such that  $1/C_3^{12}$  bounds the heights of all the the forms in  $F_0L_U^{(i)}$  from above. Then we have the following lemma estimating the number of irreducible lattice points in  $B(n, t, \lambda, X, \delta)$ .

**Proposition 4.6** *The number of irreducible integral binary cubic forms  $g$  in  $B(n, t, \lambda, X, \delta)$  that satisfy  $0 < a(g) < X^\delta$  is*

$$\begin{cases} 0 & \text{if } \frac{C_1\lambda}{t} < 1 \text{ or } \frac{C_2^{1/3}\lambda}{t} > X^{\delta/3} \text{ or } C_3\lambda < 1; \\ \text{Vol}(B(n, t, \lambda, X, \delta)) + O(\max\{\lambda^9 t^3, \lambda^9 t^{-3}, 1\}) & \text{otherwise.} \end{cases}$$

**Proof:** If  $C_1\lambda/t < 1$  or  $C_2^{1/3}\lambda/t > X^{\delta/3}$ , then no integral binary cubic form in  $B(n, t, \lambda, X, \delta)$  satisfies  $0 < a < X^\delta$ . If  $\lambda < C_3^{12}$ , then the heights of all the forms in  $B(n, t, \lambda, X, \delta)$  are less than 1 and  $B(n, t, \lambda, X, \delta)$  contains no irreducible integral forms. Otherwise, one sees that the projections of  $B(n, t, \lambda, X, \delta)$  onto  $a = 0$  and  $d = 0$  have volume  $O(\lambda^9 t^3)$  and  $O(\lambda^9 t^{-3})$ , respectively, and all other projections are also bounded by  $O(\lambda^9 t^3 + \lambda^9 t^{-3})$ . The result now follows from Proposition 2.5.  $\square$

Thus, up to an error of  $X^{3/4+\epsilon}$  due to the estimates on reducible forms in Lemma 4.5, we have

$$N(U_{\mathbb{Z}}^{(i)}; X, \delta) = \frac{1}{C_{F_0}} \int_{\lambda=C_3}^{X^{1/12}} \int_{C_2\lambda/X^{\delta/3}}^{C_1\lambda} \int_{N'} (\text{Vol}(B(n, t, \lambda, X, \delta)) + O(\max\{t^3\lambda^9, t^{-3}\lambda^9, 1\})) t^{-2} dn d^\times t d^\times \lambda. \quad (45)$$

The integral of the second summand is  $O(X^{5/6})$ , while the integral of the first summand is  $\text{Vol}(\mathcal{R}_X(L_U^{(i)}; \delta))$  up to an error of  $O(1)$ . Thus, to complete the proof of Proposition 4.3, it remains only to compute the volume of  $\mathcal{R}_X(L_U^{(i)}; \delta)$ .

#### 4.1.5 Computing the volume

The product map  $N \times A_+ \times \Lambda \rightarrow F_{\mathbb{R}}$  is a bijection and this decomposition gives a Haar measure  $dg$  on  $F_{\mathbb{R}}$ , namely,  $dg = t^{-2} dn d^\times t d^\times \lambda = t^{-2} du d^\times t d^\times \lambda$ . Write  $L_U^{(i)}$  as a union  $L_U^{(i)} = L_U^{(i),I} \cup L_U^{(i),J}$ , where  $L_U^{(i),I}$  contains the points where  $J = \pm 1$ , and  $L_U^{(i),J}$  the points where  $I = \pm 1$ . These subsets intersect in only finitely many points. We use the natural measures  $dI$  and  $dJ$  on  $L_U^{(i),I}$  and  $L_U^{(i),J}$ , respectively, and the usual Euclidean measure  $dv$  on  $U_{\mathbb{R}}$ . Then we have the following proposition.

**Proposition 4.7** *Let  $f$  be any measurable function on  $U_{\mathbb{R}}$ . Then we have*

$$(a) \quad \frac{4}{3} \int_{f_I \in L_U^{(i),I}} \int_{g \in F_{\mathbb{R}}} (\det g)^{10} f(g \cdot f_I) dg dI = \int_{v \in F_{\mathbb{R}} \cdot L_U^{(i),I}} f(v) dv;$$

$$(b) \quad \frac{4}{9} \int_{f_J \in L_U^{(i),J}} \int_{g \in F_{\mathbb{R}}} (\det g)^{10} f(g \cdot f_J) dg dJ = \int_{v \in F_{\mathbb{R}} \cdot L_U^{(i),J}} f(v) dv.$$

A simple Jacobian computation proves Proposition 4.7.

We now compute the volume of  $\mathcal{R}_X(L_U^{(i)}; \delta)$ . If  $\gamma \in F_{\mathbb{R}}$  such that  $|\det(\gamma)| = \lambda^2$ , then  $H(\gamma \cdot v) = \lambda^{12}$  for each  $v \in L_U^{(i)}$ . Therefore, by Proposition 4.7, we see that

$$\text{Vol}(\mathcal{R}_X(L_U^{(0)}; \delta)) = \frac{4}{9} \int_{J=-2}^2 \int_{\lambda=0}^{X^{1/12}} \int_{t=\lambda/X^{\delta/3}}^{\lambda} \lambda^{12} t^{-2} d^{\times} t d^{\times} \lambda dJ = \frac{4}{45} X^{5/6+2\delta/3} - \frac{4}{45} X^{5/6}. \quad (46)$$

Similarly, the volume of  $\mathcal{R}_X(L_U^{(1)}; \delta)$  is equal to

$$\frac{4}{9} \int_{J=-2}^2 \int_{\lambda=0}^{X^{1/12}} \int_{t=\lambda/X^{\delta/3}}^{\lambda} \lambda^{12} t^{-2} d^{\times} t d^{\times} \lambda dJ + 2 \cdot \frac{4}{3} \int_{I=-1}^1 \int_{\lambda=0}^{X^{1/12}} \int_{t=\lambda/X^{\delta/3}}^{\lambda} \lambda^{12} t^{-2} d^{\times} t d^{\times} \lambda dI \quad (47)$$

yielding

$$\text{Vol}(\mathcal{R}_X(L_U^{(1)}; \delta)) = \frac{4}{45} X^{5/6+2\delta/3} + \frac{4}{15} X^{5/6+2\delta/3} - \frac{16}{45} X^{5/6} = \frac{16}{45} X^{5/6+2\delta/3} - \frac{16}{45} X^{5/6}. \quad (48)$$

This concludes the proof of Proposition 4.3.

## 4.2 The number of quartic rings having submonogenized cubic resolvent rings of bounded height and bounded index

Let us now consider pairs  $(Q, (C, x))$ , where  $Q$  is a quartic ring over  $\mathbb{Z}$  and  $(C, x)$  is an submonogenized cubic resolvent ring of  $Q$  of index  $n$ . We define the *height* of  $(Q, (C, x))$  by  $H(Q, (C, x)) := H(C, x)$  and the *index* of  $(Q, (C, x))$  by  $a(Q, (C, x)) := n$ . In this subsection, we count the number of isomorphism classes of such pairs  $(Q, (C, x))$ , where  $(C, x)$  has bounded height and index. More precisely, we prove the following theorem:

**Theorem 4.8** *Let  $\delta \leq 1/4$  be a positive constant. Let  $N_4^{(i)}(X, \delta)$  denote the number of isomorphism classes of pairs  $(Q, (C, x))$ , where  $Q$  is an order in an  $S_4$ - or  $A_4$ -quartic field having  $4 - 2i$  real embeddings and  $(C, x)$  is a submonogenized cubic resolvent ring of  $Q$  having height less than  $X$  and index less than  $X^{\delta}$ . Then*

$$(a) \quad N_4^{(0)}(X, \delta) = \frac{1}{45} X^{5/6+2\delta/3} + o(X^{5/6+2\delta/3});$$

$$(b) \quad N_4^{(1)}(X, \delta) = \frac{8}{45} X^{5/6+2\delta/3} + o(X^{5/6+2\delta/3});$$

$$(c) \quad N_4^{(2)}(X, \delta) = \frac{3}{45} X^{5/6+2\delta/3} + o(X^{5/6+2\delta/3}).$$

### 4.2.1 The parametrization of quartic rings having submonogenized cubic resolvent rings of index $n$

As in §3.2, we use  $W_{\mathbb{Z}}$  to denote the space of pairs of integral ternary quadratic forms. Recall that the  $\text{GL}_2(\mathbb{Z}) \times \text{SL}_3(\mathbb{Z})$ -orbits on  $W_{\mathbb{Z}}$  are in bijective correspondence with isomorphism classes of pairs  $(Q, C)$ , where  $Q$  is a quartic ring over  $\mathbb{Z}$  and  $C$  is a cubic resolvent ring of  $Q$ .

Let  $W_{\mathbb{Z},n}$  denote the set of all  $(A, B) \in W_{\mathbb{Z}}$  such that  $4 \cdot \det(A) = n$ , and let  $H_{\mathbb{Z},1} = F_{\mathbb{Z},1} \times \text{SL}_3(\mathbb{Z})$ . We have the following proposition which follows from [4] and Proposition 4.2:

**Proposition 4.9** *The  $H_{\mathbb{Z},1}$ -orbits on  $W_{\mathbb{Z},n}$  are in canonical bijection with pairs  $(Q, (C, x))$ , where  $Q$  is a quartic ring over  $\mathbb{Z}$  and  $(C, x)$  is a submonogenized cubic resolvent of  $Q$  of index  $n$ .*

As in [5], we say that a pair  $(A, B) \in W_{\mathbb{Z}}$  is *absolutely irreducible* if  $A$  and  $B$  do not possess a common rational zero (as conics in  $\mathbb{P}^2$ ) and the cubic resolvent form  $g$  of  $(A, B)$  is irreducible over  $\mathbb{Q}$ . Equivalently, absolutely irreducible elements in  $W_{\mathbb{Z}}$  correspond to triples  $(Q, (C, x))$ , where  $Q$  is an order in an  $S_4$ - or  $A_4$ -quartic field.

Now consider the space  $W_{\mathbb{R}}$  of pairs of ternary quadratic forms over  $\mathbb{R}$ . The groups  $H_{\mathbb{R}} = F_{\mathbb{R}} \times \mathrm{SL}_3(\mathbb{R})$  and  $H_{\mathbb{R},1} = F_{\mathbb{R},1} \times \mathrm{SL}_3(\mathbb{R})$  are subgroups of  $G_{\mathbb{R}} = \mathrm{GL}_2(\mathbb{R}) \times \mathrm{SL}_3(\mathbb{R})$  and thus act on  $W_{\mathbb{R}}$  by restriction. If  $(A, B)$  is an element of  $W_{\mathbb{R}}$  and  $g(x, y) = 4 \cdot \det(Ax - By)$  is its cubic resolvent form, then the quantities

$$\begin{aligned} I(A, B) &:= I(g), \\ J(A, B) &:= J(g), \\ a(A, B) &:= a(g) \end{aligned} \tag{49}$$

are invariants for the action of  $H_{\mathbb{R},1}$  on  $W_{\mathbb{R}}$ . We define the *height* of  $(A, B)$  by

$$H(A, B) := H(g) = \max(|I|^3, J^2/4).$$

If  $(Q, (C, x))$  is the pair corresponding to the  $H_{\mathbb{Z},1}$ -equivalence class of  $(A, B) \in W_{\mathbb{Z},n}$ , then we define the invariants  $I, J$ , and  $a$  of  $(Q, (C, x))$  by  $I(Q, (C, x)) := I(A, B)$ ,  $J(Q, (C, x)) := J(A, B)$ , and  $a(Q, (C, x)) := a(A, B) = n$ . We then have  $H(Q, (C, x)) := H(C, x) = H(A, B)$ .

For  $i = 0, 1$ , and  $2$ , let  $W_{\mathbb{R},+}^{(i)}$  (resp.  $W_{\mathbb{Z},+}^{(i)}$ ) denote the set of elements  $(A, B)$  in  $W_{\mathbb{R}}$  (resp.  $W_{\mathbb{Z}}$ ) such that  $\det(A) > 0$  and  $A$  and  $B$  have  $4 - 2i$  common real zeros in  $\mathbb{P}^2$ . In the remainder of this section we prove the following result, which by Proposition 4.9 implies Theorem 4.8.

**Proposition 4.10** *For an  $H_{\mathbb{Z},1}$ -invariant set  $S \subset W_{\mathbb{Z}}$ , let  $N(S; X, \delta)$  denote the number of  $H_{\mathbb{Z},1}$ -equivalence classes of absolutely irreducible elements  $(A, B) \in S$  such that  $H(A, B) < X$  and  $0 < a(A, B) < X^{\delta}$ . Then for any positive constant  $\delta \leq 1/4$ , we have*

$$\begin{aligned} \text{(a)} \quad N(W_{\mathbb{Z},+}^{(0)}; X, \delta) &= \frac{\zeta(2)\zeta(3)}{45} X^{5/6+2\delta/3} + o(X^{5/6+2\delta/3}); \\ \text{(b)} \quad N(W_{\mathbb{Z},+}^{(1)}; X, \delta) &= \frac{8\zeta(2)\zeta(3)}{45} X^{5/6+2\delta/3} + o(X^{5/6+2\delta/3}); \\ \text{(c)} \quad N(W_{\mathbb{Z},+}^{(2)}; X, \delta) &= \frac{3\zeta(2)\zeta(3)}{45} X^{5/6+2\delta/3} + o(X^{5/6+2\delta/3}). \end{aligned}$$

## 4.2.2 Reduction theory

We first develop the necessary reduction theory in order to establish fundamental domains for the action of  $H_{\mathbb{Z},1}$  on  $W_{\mathbb{R},+}$ . We start by describing a fundamental set for the action of  $H_{\mathbb{R},1}$  on  $W_{\mathbb{R},+}$ .

First, for  $i = 0, 1$ , and  $2$ , let  $W_{\mathbb{R},+}^{(i)} := W_{\mathbb{R},[-1]}^{(i)} \cup W_{\mathbb{R},[3]}^{(i)}$ , where  $W_{\mathbb{R},[j]}^{(i)}$  denotes the set of elements  $(A, B)$  in  $W_{\mathbb{R},+}^{(i)}$  such that the signature of  $A$  is equal to  $j$ . Let  $W_{\mathbb{R},[j]}$  denote the union  $\cup_i W_{\mathbb{R},[j]}^{(i)}$ . (Note that every ternary quadratic form  $A$  over  $\mathbb{R}$  that has signature  $-1$  is equivalent to the form  $A_1$ , where  $A_1$  again denotes the first matrix in the pair of matrices in (29).)

The results of §3.2, together with the construction in §2.1 of fundamental sets for the action of  $\mathrm{GL}_2(\mathbb{R})$  on  $V_{\mathbb{R}}^{(i)}$ , imply that for  $i = 0, 1, 2+$ , and  $2-$ , the set  $L_W^{(i)}$  is a fundamental set for the action of  $H_{\mathbb{R}}$  on  $W_{\mathbb{R},[-1]}^{(i)}$ , where

$$L_W^{(i)} := \left\{ \left[ \begin{pmatrix} & & 1/2 \\ & -1 & \\ 1/2 & & \end{pmatrix}, \begin{pmatrix} a & b/2 & 0 \\ b/2 & c & d/2 \\ 0 & d/2 & e \end{pmatrix} \right] : ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4 \in L_V^{(i)} \right\}.$$

We now describe a fundamental set  $L_W^{(2\#)}$  for the action of  $H_{\mathbb{R}}$  on  $W_{\mathbb{R},[3]}$ . If  $(A, B) \in W_{\mathbb{R},[3]}$ , then the signature of  $A$  is equal to  $3$  and  $A$  is a positive definite ternary quadratic form. In particular,  $A$  and

$B$  do not have any common real zero in  $\mathbb{P}^2$  and so  $W_{\mathbb{R},[3]} \subset W_{\mathbb{R},+}^{(2)}$ . As  $A$  is positive definite, there exists  $\gamma \in H_{\mathbb{R}}$  such that  $\gamma \cdot (A, B) = (\text{id}, B')$ . Furthermore, using the spectral theorem for real symmetric matrices, we may assume that  $\gamma$  is chosen so that  $B'$  is a diagonal matrix. It follows that if  $(A, B)$  is a point in  $W_{\mathbb{R},[3]}$ , then there exists a point in  $W_{\mathbb{R},[3]}$  that is  $H_{\mathbb{R}}$ -equivalent to  $(A, B)$  and has height equal to 1. Thus, a fundamental set  $L_W^{(2\#)}$  may be constructed by choosing one point  $(A, B)$  for each  $(I, J)$  such that  $H(I, J) = 1$  and  $4I^3 - J^2 > 0$ . We choose

$$L_W^{(2\#)} := \left\{ \frac{1}{4^{1/3}} \left[ \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}, \begin{pmatrix} r_1^{(J)} & & \\ & r_2^{(J)} & \\ & & r_3^{(J)} \end{pmatrix} \right] : -2 \leq J \leq 2 \right\}$$

for our fundamental set for the action of  $H_{\mathbb{R}}$  on  $W_{\mathbb{R},[3]}$ , where the  $r_i^{(J)}$  are the three real roots of the cubic equation  $x^3 - \frac{1}{3}x - \frac{J}{27} = 0$ .

Hence  $L_W^{(0)}$ ,  $L_W^{(1)}$ , and  $L_W^{(2)} := L_W^{(2+)} \cup L_W^{(2-)} \cup L_W^{(2\#)}$  form fundamental sets for the action of  $H_{\mathbb{R}}$  on  $W_{\mathbb{R},+}^{(0)}$ ,  $W_{\mathbb{R},+}^{(1)}$  and  $W_{\mathbb{R},+}^{(2)}$ , respectively. As usual, the key fact about the  $L_W^{(i)}$  that we will need is that the absolute values of all the entries in the pair of matrices in the  $L_W^{(i)}$  are uniformly bounded.

We also require the following lemma.

**Lemma 4.11** *Let  $(A, B)$  be any point in  $W_{\mathbb{R}}^{(i)}$  having nonzero discriminant. Then the order of the stabilizer of  $(A, B)$  in  $H_{\mathbb{R}}$  is 4 if  $i = 0$  or 2, and 2 if  $i = 1$ .*

**Proof:** When  $i = 0, 1, 2+$ , or  $2-$ , the result follows from Lemma 2.2. It is now enough to check the result for points in  $L_W^{(2\#)}$ , where it is elementary.  $\square$

We now construct a fundamental domain  $\mathcal{F}_2$  for the action of  $H_{\mathbb{Z},1}$  on  $H_{\mathbb{R}}$ . We choose  $\mathcal{F}_2$  to be contained in the Siegel set  $N'A'K\Lambda$ , where

$$\begin{aligned} K_1 &= \{\text{orthogonal transformations in } H_{\mathbb{R}}\}; \\ A' &= \{(t, s_1, s_2) : t > 0, s_2, s_2 > c_1\}, \\ &\quad \text{where, } (t, s_1, s_2) = \left( \begin{pmatrix} t^{-1} & & \\ & t & \\ & & 1 \end{pmatrix}, \begin{pmatrix} s_1^{-2}s_2^{-1} & & \\ & s_1s_2^{-1} & \\ & & s_1s_2^2 \end{pmatrix} \right); \\ N' &= \{n(u_1, u_2, u_3, u_4) : |u_1|, |u_2|, |u_3|, |u_4| \leq c_2\}, \\ &\quad \text{where, } n(u_1, u_2, u_3, u_4) = \left( \begin{pmatrix} 1 & & \\ u_1 & 1 & \\ & & 1 \end{pmatrix}, \begin{pmatrix} 1 & & \\ u_2 & 1 & \\ u_3 & u_4 & 1 \end{pmatrix} \right); \\ \Lambda &= \{\lambda : \lambda > 0\}, \\ &\quad \text{where, } \lambda = \left( \begin{pmatrix} \lambda & & \\ & \lambda & \\ & & 1 \end{pmatrix}, \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \right); \end{aligned}$$

and  $c_1, c_2$  are absolute constants. We write  $\mathcal{F}_2 = \mathcal{F}_1 \times \mathcal{F}_{\text{SL}_3} \subset \text{GL}_2(\mathbb{R}) \times \text{SL}_3(\mathbb{R})$ .

Let  $n_i$  denote the size of the stabilizer in  $H_{\mathbb{R}}$  of an absolutely irreducible point in  $W_{\mathbb{R}}^{(i)}$  for  $i = 0, 1, 2$ ; by Lemma 4.11, thus  $n_0 = 4, n_1 = 2, n_2 = 4$ . Then, for any fixed  $i$ , the multiset  $\mathcal{F}_2 L_W^{(i)}$  is essentially the union of  $n_i$  fundamental domains for the action of  $H_{\mathbb{Z},1}$  on  $W_{\mathbb{R}}^{(i)}$ . More precisely, an  $H_{\mathbb{Z},1}$ -equivalence  $x$  is represented in this multiset  $m(x)$  times, where  $m(x) = \#\text{Stab}_{H_{\mathbb{R}}}(x) / \#\text{Stab}_{H_{\mathbb{Z},1}}(x)$ . Using the methods of the proof of Lemma 2.4, it is easy to see that the stabilizer in  $H_{\mathbb{Z},1}$  of an absolutely irreducible element is always trivial. We conclude that the product  $n_i \cdot N(W_{\mathbb{Z}}^{(i)}; X, \delta)$  is exactly equal to the number of absolutely irreducible integer points  $p \in \mathcal{F}_2 L_W^{(i)}$  for which  $H(p) < X$  and  $0 < a(p) < X^\delta$ .

### 4.2.3 Estimates on reducibility

By our construction of the  $L_W^{(i)}$ , the resolvent binary cubic forms of points in  $L_W^{(i)}$  lie in  $L_U^{(0)}$  for  $i = 0, 2$ , and in  $L_U^{(1)}$  for  $i = 1$ . As in §4.1.3, we consider the set  $\mathcal{R}_X(L_W^{(i)}; \delta) := \mathcal{F}_2^{(\delta)} L_W^{(i)} \cap \{p : |H(p)| < X\}$ , where  $\mathcal{F}_2^{(\delta)} \subset \mathcal{F}_2$  consists of those elements  $(\gamma_1, \gamma_2) \in \mathcal{F}_2$  that satisfy  $\gamma_1 \in \mathcal{F}_1^{(\delta)}$ . We shall need the following lemma whose proof is deferred to §4.6.

**Lemma 4.12** *Let  $\delta \leq 1/4$  be a positive constant and let  $g$  be any element in a compact subset  $H_0$  of  $H_{\mathbb{R}}$ . The number of points  $(A, B) \in \mathcal{F}_2^{(\delta)} g L_W^{(i)} \cap W_{\mathbb{Z}}$  that are not absolutely irreducible and satisfy  $a_{11} \neq 0$ ,  $H(A, B) < X$ , and  $0 < a(A, B) < X^\delta$  is  $o(X^{5/6 + 2/3\delta})$ , where the implied constant depends only on  $H_0$ .*

### 4.2.4 Averaging and cutting off the cusp

If  $\gamma = (\gamma_1, \gamma_2)$  is an element of  $H_{\mathbb{R}}$ , then we define the determinant of  $\gamma$  by  $\det(\gamma) := \det(\gamma_1)$ . Let  $H_0$  be a  $K_1$ -invariant compact set in  $H_{\mathbb{R}}$  that is the closure of some nonempty open set, such that every element in  $H_0$  has determinant greater than 1. Let  $S \subset W_{\mathbb{Z}}^{(i)}$  be an  $H_{\mathbb{Z},1}$ -invariant set and let  $S^{\text{irr}}$  denote the subset of absolutely irreducible elements in  $S$ . Then we have

$$N(S; X, \delta) = \frac{\int_{h \in H_0} \#\{x \in S^{\text{irr}} \cap \mathcal{F}_2 h L_W^{(i)} : H(x) < X, 0 < a(x) < X^\delta\} dh}{n_i \cdot \int_{h \in H_0} dh}.$$

Just as in §2.3, we see that if  $C_{H_0}^{(i)} = n_i \cdot \int_{h \in H_0} dh$ , then we have

$$N(S; X, \delta) = \frac{1}{C_{H_0}^{(i)}} \int_{g \in \mathcal{F}_2} \#\{x \in S^{\text{irr}} \cap g H_0 L_W^{(i)} : H(x) < X, 0 < a(x) < X^\delta\} dg. \quad (50)$$

Since  $H_0$  is  $K_1$ -invariant, we see that the right hand side in (50) is equal to

$$\frac{1}{C_{H_0}^{(i)}} \int_{g \in \mathcal{F}_2} \#\left\{x \in S^{\text{irr}} \cap n(t, s_1, s_2) \lambda H_0 L_W^{(i)} : H(x) < X, 0 < a(x) < X^\delta\right\} t^{-2} s_1^{-6} s_2^{-6} dn d^\times t d^\times s d^\times \lambda. \quad (51)$$

We define  $B(n, (t, s_1, s_2), \lambda, X, \delta)$  by

$$B(n, (t, s_1, s_2), \lambda, X, \delta) := \left\{x \in n(t, s_1, s_2) \lambda H_0 L_W^{(i)} : H(x) < X, 0 < a(x) < X^\delta\right\}. \quad (52)$$

Then

$$N(S; X, \delta) = \frac{1}{C_{H_0}^{(i)}} \int_{g \in \mathcal{F}_2} \#\left\{x \in S^{\text{irr}} \cap B(n, (t, s_1, s_2), \lambda, X, \delta)\right\} t^{-2} s_1^{-6} s_2^{-6} dn d^\times t d^\times s d^\times \lambda.$$

Let  $C_1$  and  $C_2$  be constants that bound the value of  $a(A, B)$  for all  $(A, B) \in H_0 L_W^{(i)}$  from above and below, respectively. Let  $C_3$  be a constant such that  $1/C_3^{12}$  bounds the height of all the the points in  $H_0 L_W^{(i)}$  from above. We then have the following proposition whose proof is identical to that of Proposition 4.6.

**Proposition 4.13** *If either  $\frac{C_1 \lambda}{t} < 1$ ,  $\frac{C_2^{1/3} \lambda}{t} > X^{\delta/3}$ , or  $C_3 \lambda < 1$ , then there are no absolutely irreducible points  $x$  in  $W_{\mathbb{Z}} \cap B(n, (t, s_1, s_2), \lambda, X, \delta)$ .*

We also have the following lemma whose proof we omit as it exactly follows that of [5, Lemma 11].

**Lemma 4.14** *Let  $h$  be an element of  $H_0$ . Then the number of absolutely irreducible elements  $(A, B) \in \mathcal{F}_2^{(\delta)} h L \cap W_{\mathbb{Z}}$  such that  $a_{11} = 0$  and  $H(A, B) < X$  is  $O(X^{5/6 + \delta/3})$ .*

Thus, to prove Proposition 4.10, we may assume that  $a_{11} \neq 0$ . If  $B(n, a(t, s_1, s_2), \lambda, X, \delta)$  contains an absolutely irreducible point  $(A, B) \in W_{\mathbb{Z}}$  with  $a_{11} \neq 0$ , then  $|\frac{ts_1^4 s_2^2}{\lambda}| = O(1)$ . From Proposition 4.13, we see that  $s_1^4 = O(X^{\delta/3})$  and  $s_2^2 = O(X^{\delta/3})$ . Equation (51) and Proposition 2.5 now imply that up to an error of  $o(X^{5/6+2\delta/3})$  coming from Lemma 4.12,  $N(W_{\mathbb{Z}}^{(i)}; X, \delta)$  is equal to

$$\frac{1}{C_{H_0}^{(i)}} \int_{\lambda=C}^{X^{1/12}} \int_{t=\frac{C_2^{1/3}\lambda}{X^{\delta/3}}}^{C_1\lambda} \int_{s_1=1/2}^{k_1 X^{\delta/12}} \int_{s_2=1/2}^{k_2 X^{\delta/6}} \int_{N'} (\text{Vol}(B(n, (t, s_1, s_2), \lambda, X, \delta)) + O(\lambda^{11} t s_1^4 s_2^3)) s_1^{-6} s_2^{-6} t^{-2} dnd^\times td^\times sd^\times \lambda, \quad (53)$$

where  $k_1$  and  $k_2$  are absolute constants. The integral of the error term in (53) is evaluated to be  $O(X^{5/6+\delta/3})$  while that of the main term is equal to

$$\begin{aligned} & \text{Vol}(\mathcal{R}_X(L_W^{(i)}; \delta)) - \frac{1}{C_{H_0}^{(i)}} \int_{\lambda=C}^{X^{1/12}} \int_{t=\frac{C_2^{1/3}\lambda}{X^{\delta/3}}}^{C_1\lambda} \int_{k_1 X^{\delta/12}}^{\infty} \int_{k_2 X^{\delta/6}}^{\infty} \int_{N'} \lambda^{12} s_1^{-6} s_2^{-6} t^{-2} dnd^\times td^\times sd^\times \lambda \\ &= \text{Vol}(\mathcal{R}_X(L_W^{(i)}; \delta)) + O(X^{5/6}). \end{aligned}$$

Therefore, to prove Proposition 4.10 it remains to compute the volume of  $\mathcal{R}_X(L_W^{(i)}; \delta)$ .

#### 4.2.5 Computing the volume

Define the usual subgroups  $\bar{N}$ ,  $A$ ,  $N$ , and  $\Lambda$  of  $H_{\mathbb{R}}$  as follows:

$$\bar{N} = \{\bar{n}(x) : x \in \mathbb{R}^4\}, \text{ where } \bar{n}(x) = \left( \begin{pmatrix} 1 & x_1 \\ & 1 \end{pmatrix}, \begin{pmatrix} 1 & x_2 & x_3 \\ & 1 & x_4 \\ & & 1 \end{pmatrix} \right); \quad (54)$$

$$A = \{(t, t_1, t_2) : t \in \mathbb{R}_+^{\times 4}\}, \text{ where } a(t) = \left( \begin{pmatrix} t^{-1} & & \\ & t & \\ & & 1 \end{pmatrix}, \begin{pmatrix} t_1 & & \\ & t_2 & \\ & & t_1^{-1} t_2^{-1} \end{pmatrix} \right); \quad (55)$$

$$N = \{n(u) : u \in \mathbb{R}^4\}, \text{ where } n(u) = \left( \begin{pmatrix} 1 & & \\ u_1 & 1 & \\ & & 1 \end{pmatrix}, \begin{pmatrix} 1 & & \\ u_2 & 1 & \\ u_3 & u_4 & 1 \end{pmatrix} \right); \quad (56)$$

$$\Lambda = \{\lambda \in \mathbb{R}_+^{\times}\}, \text{ where } \lambda = \left( \begin{pmatrix} \lambda & & \\ & \lambda & \\ & & 1 \end{pmatrix}, \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \right). \quad (57)$$

It is well-known that the natural product map  $\bar{N} \times A \times N \times \Lambda \rightarrow H_{\mathbb{R}}$  maps bijectively to a full measure set in  $H_{\mathbb{R}}$ . This decomposition gives a Haar measure  $dg$  on  $H_{\mathbb{R}}$ , namely,  $dg = t^{-2} du_1 d^\times t d^\times \lambda dh$ , where  $dh$  is the Haar measure on  $\text{SL}_3(\mathbb{R})$  given by  $dh = t_1^{-4} t_2^{-2} dx d^\times t_1 d^\times t_2 du$ .

Fix  $i$  to be one of 0, 1, 2+, 2-, or 2#. Let  $L := L_W^{(i)}$  and  $n := n_i$ . Write  $L$  as a union  $L = L^I \cup L^J$ , where  $L^I$  contains the points where  $J = \pm 1$  and  $L^J$  contains the points where  $I = \pm 1$ . These subsets again intersect only in finitely many points. Let  $W^I$  and  $W^J$  denote the sets  $H_{\mathbb{R}} \cdot L^I$  and  $H_{\mathbb{R}} \cdot L^J$ , respectively.

Let  $dI$  and  $dJ$  be the measures on  $L^I$  and  $L^J$ , respectively, and let  $dv$  be the usual Euclidean volume on  $W_{\mathbb{R}}$ . Then we have the following proposition whose proof is identical to that of Proposition 2.7.

**Proposition 4.15** *Let  $f$  be any measurable function on  $W_{\mathbb{R}}$ . Then*

- (a)  $\frac{4}{3n} \int_{(A,B) \in L^I} \int_{h \in H_{\mathbb{R}}} (\det h)^{10} f(h \cdot (A, B)) dh dI = \int_{W^I} f(v) dv;$
- (b)  $\frac{4}{9n} \int_{(A,B) \in L^J} \int_{h \in H_{\mathbb{R}}} (\det h)^{10} f(h \cdot (A, B)) dh dJ = \int_{W^J} f(v) dv.$

We now compute the volume of  $\mathcal{R}_X(L; \delta)$ . First assume that  $L$  is  $L_W^{(0)}$ ,  $L_W^{(2+)}$ ,  $L_W^{(2-)}$ , or  $L_W^{(2\#)}$ . If  $\gamma \in H_{\mathbb{R}}$  has determinant equal to  $\lambda^2$ , then  $H(\gamma \cdot v) = \lambda^{12}$  for  $v \in L$ . Therefore, by Proposition 4.15, we have

$$\text{Vol}(\mathcal{R}_X(L; \delta)) = \frac{4}{9}\zeta(2)\zeta(3) \int_{J=-2}^2 \int_{\lambda=0}^{X^{1/12}} \int_{t=\lambda/X^{\delta/3}}^{\lambda} \lambda^{12} t^{-2} d^\times t d^\times \lambda dJ = \frac{4\zeta(2)\zeta(3)}{45} (X^{5/6+2\delta/3} - X^{5/6}).$$

Similarly, if  $L = L_W^{(1)}$ , then by Proposition 4.15 the volume of  $\mathcal{R}_X(L; \delta)$  is equal to

$$\frac{4}{9}\zeta(2)\zeta(3) \int_{J=-2}^2 \int_{\lambda=0}^{X^{1/12}} \int_{t=\lambda/X^{\delta/3}}^{\lambda} \lambda^{12} t^{-2} d^\times t d^\times \lambda dJ + 2 \cdot \frac{4}{3}\zeta(2)\zeta(3) \int_{I=-1}^1 \int_{\lambda=0}^{X^{1/12}} \int_{t=\lambda/X^{\delta/3}}^{\lambda} \lambda^{12} t^{-2} d^\times t d^\times \lambda dI$$

yielding

$$\text{Vol}(\mathcal{R}_X(L_W^{(1)}; \delta)) = \left( \frac{4\zeta(2)\zeta(3)}{45} + \frac{4\zeta(2)\zeta(3)}{15} \right) (X^{5/6+2\delta/3} - X^{5/6}) = \frac{16\zeta(2)\zeta(3)}{45} (X^{5/6+2\delta/3} - X^{5/6}).$$

Since  $n_i = 4$  for  $i = 0, 2+, 2-, 2\#$  and  $n_1 = 2$ , we obtain Proposition 4.10.

### 4.3 Congruence conditions

For any set  $S$  in  $U_{\mathbb{Z}}$  (resp.  $W_{\mathbb{Z}}$ ) that is definable by congruence conditions, we denote the  $p$ -adic density of the  $p$ -adic closure of  $S$  in  $U_{\mathbb{Z}_p}$  (resp.  $W_{\mathbb{Z}_p}$ ) by  $\mu_p(S)$ , where we normalize the additive measure  $\mu_p$  on  $U_{\mathbb{Z}_p}$  (resp.  $W_{\mathbb{Z}_p}$ ) so that  $\mu_p(U_{\mathbb{Z}_p}) = 1$  (resp.  $\mu_p(W_{\mathbb{Z}_p}) = 1$ ).

We then have the following propositions whose proofs are identical to that of Theorem 2.11:

**Proposition 4.16** *Suppose  $S$  is a subset of  $U_{\mathbb{Z}}^{(i)}$  defined by finitely many congruence conditions. Then*

$$N(S; X, \delta) = N(U_{\mathbb{Z}}^{(i)}; X, \delta) \prod_p \mu_p(S) + O(X^{5/6}). \quad (58)$$

**Proposition 4.17** *Suppose  $S$  is a subset of  $W_{\mathbb{Z}}^{(i)}$  defined by finitely many congruence conditions. Then*

$$N(S; X, \delta) = N(W_{\mathbb{Z}}^{(i)}; X, \delta) \prod_p \mu_p(S) + o(X^{5/6+2\delta/3}). \quad (59)$$

### 4.4 Values of $p$ -adic densities

In order to count the number of 2-torsion elements in the class groups of submonogenized cubic fields of bounded height and index, we need to count  $F_{\mathbb{Z},1}$ -equivalence classes of *maximal* elements in  $U_{\mathbb{Z}}$  of similarly bounded height and index, where an element in  $U_{\mathbb{Z}}$  is said to be maximal if the corresponding cubic ring is maximal. We also need to count  $H_{\mathbb{Z},1}$ -equivalence classes of *strongly maximal* elements in  $W_{\mathbb{Z}}$  having bounded height and index, where a strongly maximal element of  $W_{\mathbb{Z}}$  is one whose associated binary cubic resolvent form in  $U_{\mathbb{Z}}$  is maximal.

We say that an element  $g \in U_{\mathbb{Z}}$  is *maximal at  $p$*  if the corresponding cubic ring is maximal at  $p$ . Similarly, we say that an element  $(A, B) \in W_{\mathbb{Z}}$  is *strongly maximal at  $p$*  if its associated cubic resolvent ring is maximal at  $p$ . Then an element  $g \in U_{\mathbb{Z}}$  (resp.  $(A, B) \in W_{\mathbb{Z}}$ ) is maximal (resp. strongly maximal) if and only if it is maximal (resp. strongly maximal) at all  $p$ .

We denote the set of elements in  $U_{\mathbb{Z}}$  that are maximal at  $p$  by  $\mathcal{U}_p(U)$  and the set of elements in  $W_{\mathbb{Z}}$  that are strongly maximal at  $p$  by  $\mathcal{V}_p(W)$ . We then have the following lemma (see [9], [4]).

**Lemma 4.18** *We have*

$$\begin{aligned} \mu_p(\mathcal{U}_p(U)) &= (p^3 - 1)(p^2 - 1) / p^5; \\ \mu_p(\mathcal{V}_p(W)) &= (p^3 - 1)^2(p^2 - 1)^2 / p^{10}. \end{aligned}$$



## 4.5 Uniformity estimates

We require the following two uniformity estimates in order to carry out the sieve step in the proof of Theorem 1.10.

**Proposition 4.19** *Let  $\delta \leq 1/4$  be a positive constant and let  $\mathcal{W}_p(U)$  denote the set of all binary cubic forms that correspond to cubic rings not maximal at  $p$ . Then  $N(\mathcal{W}_p(U); X, \delta) = O(X^{5/6+2\delta/3}/p^{1+4\delta/3})$ , where the implied constant is independent of  $p$ .*

**Proof:** Let  $p$  be a fixed prime and assume that  $p = X^\alpha$ . We see from [9, Lemma 9] that a cubic ring  $C$  can fail to be maximal at  $p$  in two possible ways: Either  $C$  has a  $\mathbb{Z}$ -basis  $\langle 1, \omega, \theta \rangle$  such that  $C' = \mathbb{Z} + \mathbb{Z} \cdot (\omega/p) + \mathbb{Z} \cdot \theta$  forms a cubic ring or has a  $\mathbb{Z}$ -basis  $\langle 1, \omega, \theta \rangle$  such that  $C'' = \mathbb{Z} + \mathbb{Z} \cdot (\omega/p) + \mathbb{Z} \cdot (\theta/p)$  forms a cubic ring. Now  $H(C', x) = H(C, x)/p^2$  and  $H(C'', x) = H(C, x)/p^4$ . Furthermore, as  $a(C, x)$  is equal to the index of  $\mathbb{Z}[x]$  in  $C$ , it follows that  $a(C', x) = p \cdot a(C, x)$  and that  $a(C'', x) = p^2 \cdot a(C, x)$ . Thus, identically as in [9, Proposition 20], we see that

$$N(\mathcal{W}_p(U); X, \delta) = O(N(U_{\mathbb{Z}}; X^{1-2\alpha}, \delta + \alpha) + N(U_{\mathbb{Z}}; X^{1-4\alpha}, \delta + 2\alpha)).$$

The result now follows from Proposition 4.3.  $\square$

**Proposition 4.20** *Let  $\delta \leq 1/4$  be a positive constant and let  $\mathcal{W}_p(W)$  denote the set of all  $(A, B) \in W_{\mathbb{Z}}$  that are not strongly maximal at  $p$ . Then  $N(\mathcal{W}_p(W); X, \delta) = O(X^{5/6+2\delta/3}/p^{1+4\delta/3})$  where the implied constant is independent of  $p$ .*

**Proof:** Let  $\mathcal{W}_p^{(1)}(W)$  be the set of all  $(A, B) \in W_{\mathbb{Z}}$  that are nonmaximal at  $p$  and let  $\mathcal{W}_p^{(2)}(W)$  be the set of all  $(A, B) \in W_{\mathbb{Z}}$  that are maximal at  $p$  but also overramified at  $p$ . A cubic resolvent ring of a quartic ring  $Q$  is maximal if and only if  $Q$  is maximal and not overramified. Thus  $N(\mathcal{W}_p(W); X, \delta) = N(\mathcal{W}_p^{(1)}(W); X, \delta) + N(\mathcal{W}_p^{(2)}(W); X, \delta)$ .

Now suppose that  $p$  is equal to  $X^\alpha$ . Then an argument identical to the proof of [5, Proposition 23] shows that we have

$$N(\mathcal{W}_p^{(1)}(W); X, \delta) = O(N(W_{\mathbb{Z}}; X^{1-2\alpha}, \delta + \alpha)) = O(X^{(1-2\alpha)(\frac{5}{6} + \frac{2}{3}(\delta + \alpha))}) = O(X^{\frac{5}{6} + \frac{2}{3}\delta - \alpha - \frac{4}{3}\alpha\delta}),$$

$$N(\mathcal{W}_p^{(2)}(W); X, \delta) = O(N(W_{\mathbb{Z}}; X^{1-2\alpha}, \delta + \alpha)) = O(X^{(1-2\alpha)(\frac{5}{6} + \frac{2}{3}(\delta + \alpha))}) = O(X^{\frac{5}{6} + \frac{2}{3}\delta - \alpha - \frac{4}{3}\alpha\delta}),$$

yielding the proposition.  $\square$

The results of this section also allow us now to finally prove the key uniformity estimate for binary quartic forms that was used in Section 3, namely, Proposition 3.18. This will complete the proof of Theorem 1.9. We note that Proposition 3.18 will also prove a key role in the proofs of Theorems 1.1 and 1.3.

**Proof of Proposition 3.18:** Let  $\mathcal{W}_p^{(1)}(V)$  be the set of all binary quartic forms  $f \in V_{\mathbb{Z}}$  that are nonmaximal at  $p$  and let  $\mathcal{W}_p^{(2)}(V)$  be the set of all forms  $f \in V_{\mathbb{Z}}$  that are maximal at  $p$  but also overramified at  $p$ . Clearly  $N(\mathcal{W}_p(V); X) = N(\mathcal{W}_p^{(1)}(V); X) + N(\mathcal{W}_p^{(2)}(V); X)$ . Then, as in [5, Proposition 23], we have

$$N(\mathcal{W}_p^{(1)}(V); X) = O(N(W_{\mathbb{Z}}; X^{1-2\alpha}, \alpha)) = O(X^{(1-2\alpha)(\frac{5}{6} + \frac{2}{3}\alpha)}) = O(X^{\frac{5}{6} - \alpha - \frac{4}{3}\alpha^2}),$$

$$N(\mathcal{W}_p^{(2)}(V); X) = O(N(W_{\mathbb{Z}}; X^{1-2\alpha}, \alpha)) = O(X^{(1-2\alpha)(\frac{5}{6} + \frac{2}{3}\alpha)}) = O(X^{\frac{5}{6} - \alpha - \frac{4}{3}\alpha^2}).$$

Consider the pair  $(Q, (C, x))$ , where  $Q$  is a quartic ring and  $(C, x)$  is a submonogenized cubic resolvent ring of  $Q$ . The discriminant of the pair  $(Q, (C, x))$  is less than twice the height of  $(Q, (C, x))$ . Therefore, the fact that a cubic ring has at most 12 monogenizations (see [25] and [27]) combined with [5, Proposition 23] also yields the estimate

$$N(\mathcal{W}_p(V); X) = O(X/p^2) = O(X^{1-2\alpha}).$$

Combining the three estimates we obtain that  $N(\mathcal{W}_p(V); X) = O(X^{5/6}/p^{1+2/3})$ , as desired.  $\square$

## 4.6 Proof of Lemma 4.12

The proof of this lemma is very similar to that of [6, Lemma 13]. Recall that if an element  $(A, B) \in W_{\mathbb{Z}}$  corresponding to the pair of rings  $(Q, C)$  is not absolutely irreducible, then  $Q$  is not an order in an  $S_4$ -field. This implies that (4), (13), and (112) cannot all occur as splitting types of primes in  $Q$ . Indeed, if all three of these splitting types occur in  $Q$ , then  $Q$  must be a domain, and the Galois group  $G$  of the Galois closure of the quotient field of  $Q$  over  $\mathbb{Q}$  must contain a 4-cycle, a 3-cycle, and a 2-cycle, implying that  $G = S_4$ .

Let us denote by  $T_p(W, \theta)$  the set of elements in  $W_{\mathbb{Z}}$  whose splitting type over  $p$  is  $\theta$ . It follows from [4, Lemma 23] that  $\mu_p(T_p(W, (4)))$  approaches  $1/4$  as  $p \rightarrow \infty$ . Therefore the total number of points  $(A, B) \in \mathcal{F}_2^{(\delta)} hL_W^{(i)}$  having height less than  $X$ ,  $a_{11} \neq 0$ , and such that the quartic ring associated to  $(A, B)$  does not have splitting type (4) over any finite prime  $p$  is

$$O\left(\lim_{N \rightarrow \infty} \prod_{p < N} (1 - \mu_p(T_p(W, (4)))) X^{\frac{5}{6} + \frac{2\delta}{3}}\right) = o(X^{5/6 + 2\delta/3}).$$

Similarly, the total number of points  $(A, B) \in \mathcal{F}_2^{(\delta)} hL_W^{(i)}$  having height less than  $X$ ,  $a_{11} \neq 0$  in  $\mathcal{F}_2^{(\delta)} hL_W^{(i)}$ , and such that  $(A, B)$  does not have splitting type (112) (resp. (13)) is also  $o(X^{5/6 + 2\delta/3})$ . The lemma follows.

## 4.7 Proof of the main theorem (Theorem 1.10)

By Lemma 4.18, we have  $\mu(\mathcal{U}_p(U)) = (p^3 - 1)(p^2 - 1)/p^5$  and  $\mu(\mathcal{V}_p(W)) = (p^2 - 1)^2(p^3 - 1)^2/p^{10}$ . Identically as in the proof of Theorem 1.9, we use Theorems 4.1 and 4.8 and Propositions 4.19 and 4.20 to obtain

$$\begin{aligned} \lim_{X \rightarrow \infty} \frac{\sum_{\substack{|H(K)| < X \\ \text{Disc}(K) > 0}} (\#\text{Cl}_2(K) - 1)}{\sum_{\substack{|H(K)| < X \\ \text{Disc}(K) > 0}} 1} &= \frac{\frac{\zeta(2)\zeta(3)}{45} X^{5/6 + 2\delta/3}}{\frac{4}{45} X^{5/6 + 2\delta/3}} \prod \frac{\mu(\mathcal{V}_p(W))}{\mu(\mathcal{U}_p(U))} = \frac{1}{4} \frac{\zeta(2)\zeta(3)\zeta(2)\zeta(3)}{\zeta(2)^2\zeta(3)^2} = \frac{1}{4}, \\ \lim_{X \rightarrow \infty} \frac{\sum_{\substack{|H(K)| < X \\ \text{Disc}(K) < 0}} (\#\text{Cl}_2(K) - 1)}{\sum_{\substack{|H(K)| < X \\ \text{Disc}(K) < 0}} 1} &= \frac{\frac{8\zeta(2)\zeta(3)}{45} X^{5/6 + 2\delta/3}}{\frac{16}{45} X^{5/6 + 2\delta/3}} \prod \frac{\mu(\mathcal{V}_p(W))}{\mu(\mathcal{U}_p(U))} = \frac{1}{2} \frac{\zeta(2)\zeta(3)\zeta(2)\zeta(3)}{\zeta(2)^2\zeta(3)^2} = \frac{1}{2}, \\ \lim_{X \rightarrow \infty} \frac{\sum_{\substack{|H(K)| < X \\ \text{Disc}(K) > 0}} (\#\text{Cl}_2^+(K) - 1)}{\sum_{\substack{|H(K)| < X \\ \text{Disc}(K) > 0}} 1} &= \frac{\frac{4\zeta(2)\zeta(3)}{45} X^{5/6 + 2\delta/3}}{\frac{4}{45} X^{5/6 + 2\delta/3}} \prod \frac{\mu(\mathcal{V}_p(W))}{\mu(\mathcal{U}_p(U))} = \frac{\zeta(2)\zeta(3)\zeta(2)\zeta(3)}{\zeta(2)^2\zeta(3)^2} = 1, \end{aligned}$$

where  $K$  ranges over isomorphism classes of submonogenized cubic fields having index at most  $X^\delta$ . This proves the weaker version of Theorem 1.10 where we average over all submonogenized fields of index at most  $X^\delta$  and height at most  $X$ , without any specified splitting conditions.

To complete the proof of Theorem 1.10 in all cases, we proceed as in §3.7. Let  $p$  be a fixed prime. For a cubic splitting type  $\sigma$ , let  $\mathcal{U}_p(U, \sigma)$  denote the set of integral binary cubic forms in  $\mathcal{U}_p(U)$  whose associated cubic ring has splitting type  $\sigma$  over  $p$ . Similarly, for a quartic splitting type  $\theta$ , let  $\mathcal{V}_p(W, \theta)$  denote the set of elements in  $\mathcal{V}_p(W)$  whose associated quartic ring has splitting type  $\theta$  over  $p$ . Then it suffices to prove that

$$\frac{\mu_p(\mathcal{U}_p(U, \sigma))}{\mu_p(\mathcal{U}_p(U))} = \frac{\sum_{\theta \in R^{-1}(\sigma)} \mu_p(\mathcal{V}_p(W, \theta))}{\mu_p(\mathcal{V}_p(W))}$$

for all values of  $p$  and  $\sigma$ . This equality follows from the computations in [9, Lemma 10] and [4, Lemma 23], and is summarized in Table 4; we have proven Theorem 1.10.

$\sigma$	$\frac{\mu_p(\mathcal{U}_p(U, \sigma))}{\mu_p(\mathcal{U}_p(U))}$	$R^{-1}(\sigma)$	$\frac{\sum_{\theta \in R^{-1}(\sigma)} \mu_p(\mathcal{V}_p(W, \theta))}{\mu_p(\mathcal{V}_p(W))}$
(111)	$\frac{p^2}{6(p^2 + p + 1)}$	$\{(1111), (22)\}$	$\frac{p^2}{24(p^2 + p + 1)} + \frac{p^2}{8(p^2 + p + 1)} = \frac{p^2}{6(p^2 + p + 1)}$
(12)	$\frac{p^2}{2(p^2 + p + 1)}$	$\{(112), (4)\}$	$\frac{p^2}{4(p^2 + p + 1)} + \frac{p^2}{4(p^2 + p + 1)} = \frac{p^2}{2(p^2 + p + 1)}$
(3)	$\frac{p^2}{3(p^2 + p + 1)}$	$\{(13)\}$	$\frac{p^2}{3(p^2 + p + 1)}$
(1 <sup>2</sup> 1)	$\frac{p}{p^2 + p + 1}$	$\{(1^211), (1^22)\}$	$\frac{p}{2(p^2 + p + 1)} + \frac{p}{2(p^2 + p + 1)} = \frac{p}{p^2 + p + 1}$
(1 <sup>3</sup> )	$\frac{1}{p^2 + p + 1}$	$\{(1^31)\}$	$\frac{1}{p^2 + p + 1}$

Table 4: Demonstration of the equality of  $p$ -adic density ratios for  $U$  and  $W$

## 5 The mean size of the 2-Selmer group of elliptic curves

Recall that every elliptic curve  $E$  over  $\mathbb{Q}$  can be written in the form

$$E(A, B) : y^2 = x^3 + Ax + B, \quad (60)$$

where  $A, B \in \mathbb{Z}$  and  $p^4 \nmid A$  if  $p^6 \mid B$ . For any elliptic curve  $E(A, B)$  over  $\mathbb{Q}$  written in the form (60), we define the quantities  $I(E)$  and  $J(E)$  by

$$\begin{aligned} I(E) &:= -3A, \\ J(E) &:= -27B. \end{aligned}$$

In the introduction, we had defined the height of  $E(A, B)$  by

$$H(E(A, B)) = \max\{4|A^3|, 27B^2\}.$$

In this section, we shall work with the slightly different height  $H'(E)$  defined by

$$H'(E) := \max(|I(E)|^3, J(E)^2/4).$$

Note that  $H$  and  $H'$  only differ by a constant factor; namely, for every elliptic curve  $E$  over  $\mathbb{Q}$  we have  $27H(E) = 4H'(E)$ .

In this section, we prove Theorem 1.3 by computing the average size of the 2-Selmer group of rational elliptic curves when these curves are ordered by their heights (note that the two heights  $H$  and  $H'$  give the same ordering on every set of elliptic curves). Theorem 1.1, being a special case of Theorem 1.3, will then follow.

In fact, we prove a stronger statement than Theorem 1.3. To state this result, we need some notation. If  $F$  is a set of elliptic curves over  $\mathbb{Q}$  defined by congruence conditions, then we denote by  $F^{\text{inv}}$  the set  $\{(I(E), J(E)) : E \in F\}$ . We denote the  $p$ -adic closure of  $F^{\text{inv}}$  in  $\mathbb{Z}_p \times \mathbb{Z}_p$  by  $F_p^{\text{inv}}$ . We say that such a set  $F$  of elliptic curves over  $\mathbb{Q}$  is *large at  $p$*  if the set of all monogenized cubic rings over  $\mathbb{Z}_p$  having invariants equal to  $(I, J) \in F_p^{\text{inv}}$  contains all the maximal monogenized cubic rings over  $\mathbb{Z}_p$ . We say that a set of elliptic curves  $F$  is *large* if it is large at all but finitely many primes  $p$ . In this section, we prove the following strengthening of Theorem 1.3.

**Theorem 5.1** *When elliptic curves  $E$  in any large family are ordered by height, the average size of the 2-Selmer group  $S_2(E)$  is 3.*

Note that the set of all elliptic curves is large. So too is the set of elliptic curves  $E : y^2 = g(x)$  defined by finitely many congruence conditions on the coefficients of  $g$ . Thus Theorem 1.3 indeed follows from Theorem 5.1.

**Remark on notation** Unlike the introduction, we denote the elliptic curve  $E : y^2 = x^3 + Ax + B$  by  $E(A, B)$  in this section. Also, from hereon in, we denote the elliptic curve having invariants equal to  $I$  and  $J$  by  $E_{I,J}$ .

## 5.1 Integral binary quartic forms and the 2-Selmer group of elliptic curves

Recall that an element in the 2-Selmer group of an elliptic curve  $E/\mathbb{Q}$  may be thought of as a locally soluble 2-covering of  $E/\mathbb{Q}$ . Moreover, a locally soluble 2-covering of  $E/\mathbb{Q}$  that is defined over  $\mathbb{Q}$  has a degree 2 divisor defined over  $\mathbb{Q}$ , thus giving a binary quartic form over  $\mathbb{Q}$  well-defined up to  $\text{GL}_2(\mathbb{Q})$ -equivalence.

We say that two binary quartic forms  $f_1$  and  $f_2$  with coefficients in a field  $K$  are  $K$ -equivalent if there exists  $\mu \in K$  and  $\gamma \in \text{GL}_2(K)$  such that  $f_1 = \mu^2(\gamma \cdot f_2)$ . For ease of notation, if two binary quartic forms over  $\mathbb{Q}$  are  $\mathbb{Q}$ -equivalent, then we simply say that they are *equivalent*. We then have the following lemma (see [12, Lemma 2] and the discussion following it).

**Lemma 5.2** *Let  $E$  be an elliptic curve defined over  $\mathbb{Q}$ . If  $\theta : \mathcal{C} \rightarrow E$  is a locally soluble 2-covering in  $G$ , then there is a model for  $\mathcal{C}$  of the form*

$$z^2 = f(x) \equiv ax^4 + bx^3 + cx^2 + dx + e \quad (61)$$

where  $a, b, c, d, e \in \mathbb{Q}$ . Moreover, the binary quartic form  $F(x, y) = ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4$  has invariants equal to  $\lambda^4 I(E)$  and  $\lambda^6 J(E)$  for some  $\lambda \in \mathbb{Q}$ . Finally, two 2-coverings  $C_1$  and  $C_2$  of  $E$  represented by  $f_1(x)$  and  $f_2(x)$ , respectively, give isomorphic 2-coverings of  $E$  if and only if  $F_1(x, y)$  and  $F_2(x, y)$  are equivalent.

We say that a binary quartic form  $f \in V_{\mathbb{Q}}$  is *locally soluble* if the equation  $z^2 = f(x, y)$  has solutions in  $\mathbb{R}$  and in  $\mathbb{Q}_p$  for all primes  $p$ . Since every binary quartic form with rational coefficients is clearly equivalent to one with integer coefficients, Lemma 5.2 implies that if  $E/\mathbb{Q}$  is an elliptic curve having invariants  $I$  and  $J$ , then the set of elements in the 2-Selmer group of  $E$  is in bijective correspondence with the set of equivalence classes of locally soluble integral binary quartic forms having invariants  $\lambda^4 I$  and  $\lambda^6 J$  for some  $\lambda \in \mathbb{Q}$ .

We now recall the following facts which are Lemmas 3, 4, and 5 of [12].

**Lemma 5.3** *Let  $f$  be an integral binary quartic form having invariants  $I$  and  $J$ . Let  $p \geq 5$  be a prime such that  $p^4 \mid I$ ,  $p^6 \mid J$ , and the equation  $z^2 = f(x, y)$  has a solution over  $\mathbb{Q}_p$ . Then there exists an integral binary quartic form  $f'$  equivalent to  $f$  having invariants  $p^{-4}I$  and  $p^{-6}J$ .*

**Lemma 5.4** *Let  $f$  be an integral binary quartic form having invariants  $I$  and  $J$ . Suppose that  $3^5 \mid I$ ,  $3^9 \mid J$  and  $z^2 = f(x, y)$  has a solution in  $\mathbb{Q}_3$ . Then there exists an integral binary quartic form  $f'$  equivalent to  $f$  having invariants  $3^{-4}I$  and  $3^{-6}J$ .*

**Lemma 5.5** *Let  $f$  be an integral binary quartic form having invariants  $I$  and  $J$ . Suppose that  $2^6 \mid I$ ,  $2^9 \mid J$ ,  $2^{10} \mid (8I + J)$ , and  $z^2 = g(x, y)$  has a solution in  $\mathbb{Q}_2$ . Then there exists an integral binary quartic form  $f'$  equivalent to  $f$  having invariants  $2^{-4}I$  and  $2^{-6}J$ .*

We may use these lemmas to prove the following theorem.

**Theorem 5.6** *Let  $E$  be a fixed elliptic curve over  $\mathbb{Q}$  having nonzero discriminant and invariants equal to  $I$  and  $J$ . Then the elements of the 2-Selmer group of  $E$  are in one-to-one correspondence with  $\text{PGL}_2(\mathbb{Q})$ -equivalence classes of locally soluble integral binary quartic forms having invariants equal to  $2^4 I$  and  $2^6 J$ .*

**Proof:** Let  $f$  be a locally soluble integral binary quartic form having invariants equal to  $\lambda^4 I$  and  $\lambda^6 J$  for some  $\lambda \in \mathbb{Q}$ . We first prove the existence of an integral binary quartic form  $f'$  equivalent to  $f$  such that  $f'$  has invariants equal to  $2^4 I$  and  $2^6 J$ .

We find such an  $f'$  as follows. Firstly, note we may assume that  $\lambda \in \mathbb{Z}$  and that 2 divides  $\lambda$ . Indeed, if  $\lambda = p/q$  with  $p, q \in \mathbb{Z}$ , then we may replace  $f$  by  $4q^2f$ . Secondly, by Lemma 5.3, we may assume that no prime  $p \geq 5$  divides  $\lambda$ . Now if  $3^k$  divides  $\lambda$  for some integer  $k \geq 1$ , then since  $I = I(E)$  is divisible by 3 and  $J = J(E)$  is divisible by 27, we see that  $3^5$  divides  $I(f) = \lambda^4 I$  and  $3^9$  divides  $J(g) = \lambda^6 J$ . Therefore, we may repeatedly use Lemma 5.4 to assure that 3 does not divide  $\lambda$ . Finally, if 4 divides  $\lambda$ , then  $2^8$  divides  $I$  and  $2^{12}$  divides  $J$ . Now, repeated use of Lemma 5.5 yields an integral binary quartic form  $f'$ , having invariants equal to  $2^4 I$  and  $2^6 J$ , that is equivalent to  $f$ .

Therefore, by Lemma 5.2, the set of elements in the 2-Selmer group of  $E$  are in one to one correspondence with equivalence classes of locally soluble integral binary quartic forms having invariants equal to  $2^4 I(E)$  and  $2^6 J(E)$ . Theorem 5.6 now follows from the observation that two integral binary quartic forms having the same invariants are equivalent if and only if they are  $\mathrm{PGL}_2(\mathbb{Q})$ -equivalent (if  $g = \mu^2(\gamma \cdot f)$ , then  $I(g) = \mu^4(\det \gamma)^4$  and  $J(g) = \mu^6(\det \gamma)^6$  and so  $\mu^2 = 1/(\det \gamma)^2$ . Thus  $f$  and  $g$  are  $\mathrm{PGL}_2(\mathbb{Q})$ -equivalent.)  $\square$

An elliptic curve  $E : y^2 = x^3 + Ax + B$  over  $\mathbb{Q}$  has a non-trivial 2-torsion point defined over  $\mathbb{Q}$  if and only if the corresponding cubic equation  $x^3 + Ax + B$  has a rational root. We have the following two propositions; the first follows immediately from Lemma 3.3, while the second follows from Lemma 3.21 and Theorem 5.6.

**Proposition 5.7** *The number of elliptic curves  $E$  over  $\mathbb{Q}$  such that  $E$  has a non-trivial rational 2-torsion point and  $H(E) < X$  is  $O(X^{1/2+\epsilon})$ .*

**Proposition 5.8** *The total number of elements in the union of the 2-Selmer groups of all elliptic curves over  $\mathbb{Q}$  having a nontrivial rational 2-torsion point and height bounded by  $X$  is  $O(X^{3/4+\epsilon})$ .*

We say that an elliptic curve  $E$  over a field  $K$  is *rigid* if  $I(E) \neq 0$  and  $J(E) \neq 0$ ; rigid elliptic curves over  $K$  are precisely those elliptic curves that have no extra automorphisms over  $\bar{K}$  (see e.g., [40, Thm. 10.1]). Counting just those points with  $I = 0$  or  $J = 0$  in the proofs of Proposition 2.10 and Theorem 2.1, we easily obtain the following proposition:

**Proposition 5.9** *The number of non-rigid elliptic curves over  $\mathbb{Q}$  having height bounded by  $X$  is  $O(X^{1/2})$ . The total number of elements in the union of the 2-Selmer groups of non-rigid elliptic curves having height bounded by  $X$  is  $O(X^{3/4})$ .*

We will see that the total number of elliptic curves having height bounded by  $X$  in any large family is  $\gg X^{5/6}$ . Thus, for the purposes of proving Theorem 5.1, it suffices to consider only those elliptic curves over  $\mathbb{Q}$  that are rigid and have no nontrivial rational 2-torsion point.

## 5.2 Computations of $p$ -adic densities in terms of local masses

Suppose  $K$  is a field and  $E$  is an elliptic curve defined over  $K$ . As in the special case  $K = \mathbb{Q}$  discussed in the introduction, the  $K$ -isomorphism classes of soluble 2-coverings  $C$  of  $E$  correspond bijectively to elements of  $E(K)/2E(K)$ . These soluble 2-coverings  $C$  then possess a degree two line bundle over  $K$ , thereby yielding a double cover of  $\mathbb{P}^1(K)$  ramified at four points. We thus obtain a  $K$ -soluble binary quartic form with coefficients in  $K$ , where a binary quartic form  $f$  over  $K$  is called  $K$ -soluble if the equation  $z^2 = f(x, y)$  has a solution in  $K$ . This yields the following lemma.

**Lemma 5.10** *Let  $E$  be an elliptic curve over a field  $K$  having invariants  $I \neq 0$  and  $J \neq 0$ . Then there exists a natural map*

$$\mathcal{Q}_E : E(K)/2E(K) \rightarrow \{K\text{-equivalence classes of quartics}\}.$$

*This map is injective and the image consists exactly of the  $K$ -soluble  $K$ -equivalence classes of quartics having invariants equal to  $I$  and  $J$ .*

See, e.g., [19, Proposition 2.2] for an explicit construction of  $\mathcal{Q}_E$  when  $K$  has characteristic not 2 or 3.

Now a 2-covering  $\mathcal{C}$  of an elliptic curve  $E$  over  $K$  can be viewed as a principal homogeneous space for  $E/K$ . The curves  $E$  and  $\mathcal{C}$  are isomorphic over  $\overline{K}$  and the automorphisms of  $\mathcal{C}$  over  $\overline{K}$  correspond to the 2-torsion points of  $E(\overline{K})$ . In fact, a 2-torsion point of  $E(\overline{K})$  yields an automorphism of  $\mathcal{C}$  over  $\overline{K}$  simply via the action of  $E$  on  $\mathcal{C}$ . Moreover, this automorphism is defined over  $K$  if and only if the 2-torsion point in question is itself defined over  $K$ . In conjunction with Lemma 5.10, the above discussion yields the following lemma.

**Lemma 5.11** *Let  $E$  be an elliptic curve defined over a field  $K$  having nonzero invariants. If  $\sigma \in E(K)/2E(K)$  corresponds to the  $K$ -equivalence class of the binary quartic form  $f$  under the correspondence of Lemma 5.10, then the size of the stabilizer of  $f$  in  $\mathrm{PGL}_2(K)$  is equal to  $\#E(K)[2]$ , the number of 2-torsion elements in  $E$  defined over  $K$ .*

We now turn to Theorem 5.6, which asserts that elements that are not the identity in the 2-Selmer group of an elliptic curve  $E_{I,J}$  over  $\mathbb{Q}$  are in bijective correspondence with  $\mathrm{PGL}_2(\mathbb{Q})$ -equivalence classes of locally soluble irreducible integral binary quartic forms having invariants  $2^4I$  and  $2^6J$ . In §2, we computed the number of  $\mathrm{GL}_2(\mathbb{Z})$ -equivalence classes of integral binary quartic forms having bounded height. In order to use the latter result to compute the number of  $\mathrm{PGL}_2(\mathbb{Q})$ -orbits of integral binary quartic forms, we construct a set of integral binary quartic forms by choosing, from each such  $\mathrm{PGL}_2(\mathbb{Q})$ -orbit, one  $\mathrm{GL}_2(\mathbb{Z})$ -orbit of integral binary quartic forms. In the remainder of this section, we compute the  $p$ -adic density of the union of these  $\mathrm{GL}_2(\mathbb{Z})$ -orbits.

To this end, for any  $(I, J) \in \mathbb{Z} \times \mathbb{Z}$ , let  $C_1, \dots, C_k$  be a set of  $\mathrm{PGL}_2(\mathbb{Z})$ -orbits on the set of locally soluble integral binary quartic forms having invariants equal to  $I$  and  $J$ . Let  $C_{i_1}, \dots, C_{i_l}$  be a maximal set of  $\mathrm{PGL}_2(\mathbb{Q})$ -inequivalent classes among the  $C_1, \dots, C_k$ , and define  $S^{I,J}$  to be the union of all the integral binary quartic forms in the  $C_{i_j}$  over all  $j$ . Define  $S^F$  to be the union of  $S^{2^4I, 2^6J}$  over all  $(I, J) \in F^{\mathrm{inv}}$ . We now determine the  $p$ -adic density  $\mu_p(S^F)$  of the  $p$ -adic closure of  $S^F$  in  $V_{\mathbb{Z}_p}$  in terms of a local ( $p$ -adic) mass  $M_p(V, F)$  of all isomorphism classes of soluble 2-coverings of elliptic curves over  $\mathbb{Q}_p$ :

**Proposition 5.12** *We have  $\mu_p(S^F) = |2^{10}/3^3|_p \cdot M_p(V, F)$ , where*

$$M_p(V, F) := \left(1 - \frac{1}{p^2}\right) \int_{(I,J) \in F_p^{\mathrm{inv}}} \sum_{\sigma \in E_{I,J}/2E_{I,J}} \frac{1}{\#E_{I,J}[2](\mathbb{Q}_p)} dIdJ.$$

**Proof:** Let  $F_p^{\mathrm{inv}}$  again denote the  $p$ -adic closure of  $F^{\mathrm{inv}}$  in  $\mathbb{Z}_p \times \mathbb{Z}_p$ , and let  $S_p^F$  be the  $p$ -adic closure of  $S^F$  in  $V_{\mathbb{Z}_p}$ . For  $(I, J) \in F_p^{\mathrm{inv}}$ , let  $B_p^{I,J}$  consist of a set of representatives  $f_1, \dots, f_k \in V_{\mathbb{Z}_p}$  for the action of  $\mathrm{PGL}_2(\mathbb{Q}_p)$  on the set of soluble binary quartic forms in  $V_{\mathbb{Z}_p}$  having invariants equal to  $2^4I$  and  $2^6J$ . Define  $B_p^F$  by

$$B_p^F := \bigcup_{\substack{(I,J) \in F_p^{\mathrm{inv}} \\ f \in B_p^{I,J}}} \mathrm{PGL}_2(\mathbb{Z}_p) \cdot f.$$

The  $p$ -adic density  $\mu_p(S^F)$  can be determined from the set  $B_p^F$ ; namely, we have

$$\int_{v \in S_p^F} dv = \int_{f \in B_p^F} \frac{1}{\#\mathrm{Aut}(f)} df,$$

where  $\mathrm{Aut}(f)$  denotes the stabilizer of  $f$  in  $\mathrm{PGL}_2(\mathbb{Q}_p)$ . The latter integral can be computed using a Jacobian change of variables; indeed, Proposition 2.8 and the principle of permanence of identities imply that

$$\int_{f \in B_p^F} \frac{1}{\#\mathrm{Aut}(f)} df = |2^{10} \cdot 2/27|_p \mathrm{Vol}(\mathrm{PGL}_2(\mathbb{Z}_p)) \int_{(I,J) \in F_p^{\mathrm{inv}}} \sum_{f \in B_p^{I,J}} \frac{1}{\#\mathrm{Aut}(f)}.$$

Note that we have the factor  $2^{10} \cdot 2/27$  instead of  $2/27$  because  $f \in B_p^{I,J}$  has invariants  $2^4I$  and  $2^6J$  as opposed to  $I$  and  $J$ .

By Lemmas 5.10 and 5.11, elements  $f \in B_p^{I,J}$  correspond bijectively with elements  $\sigma \in E_{I,J}/2E_{I,J}$ , and the cardinality of  $\text{Aut}(f)$  is equal to the cardinality of  $E_{I,J}(\mathbb{Q}_p)[2]$  whenever  $I \neq 0$  and  $J \neq 0$ . Proposition 5.12 now follows because the volume of  $\text{PGL}_2(\mathbb{Z}_p)$  with respect to the Haar measure obtained from the  $\bar{N}AN$  decomposition of  $\text{PGL}_2(\mathbb{Q}_p)$  is equal to  $(1 - 1/p^2)$  for  $p \geq 3$  and to  $2(1 - 1/2^2)$  for  $p = 2$  (along with the fact that the measure of the set  $\{(0, J)\} \cup \{(I, 0)\} \subset F^{\text{inv}}$  is equal to 0).  $\square$

In the analogous manner, if  $F$  is a large set of elliptic curves, then we may define  $M_p(U_1, F)$  to be the measure of  $F_p^{\text{inv}}$  with respect to the measure  $dI dJ$  on  $\mathbb{Z}_p \times \mathbb{Z}_p$ , where the measure  $dI dJ$  is normalized so that  $\mathbb{Z}_p \times \mathbb{Z}_p$  has measure 1. That is, we have

$$M_p(U_1, F) = \int_{(I,J) \in F_p^{\text{inv}}} dI dJ. \quad (62)$$

### 5.3 Uniformity estimates

In this subsection, we prove a uniformity estimate that, analogous to the proofs of Theorems 1.9 and 1.10, allows us to carry out the sieve step in the proof of Theorem 5.1.

We say that an element  $f \in V_{\mathbb{Z}}$  is *bad at  $p$*  if either  $f$  is not  $\mathbb{Q}_p$ -soluble or there exists an element  $f' \in V_{\mathbb{Z}}$  such that  $f$  is  $\text{PGL}_2(\mathbb{Q}_p)$ -equivalent to  $f'$  but not  $\text{PGL}_2(\mathbb{Z}_p)$ -equivalent to  $f'$ . Then we have the following proposition:

**Proposition 5.13** *The number of  $\text{PGL}_2(\mathbb{Z})$ -equivalence classes of elements in  $V_{\mathbb{Z}}$  that are bad at  $p$  and have height less than  $X$  is  $O(X^{5/6}/p^{5/3})$ , where the implied constant is independent of  $p$ .*

**Proof:** Let  $p \geq 5$  be a fixed prime. We show that if an element  $f \in V_{\mathbb{Z}}$  is bad at  $p$ , then  $f$  is not strongly maximal at  $p$ . Then Proposition 3.18 (proved in §4.5) will imply the result.

If  $f, f' \in V_{\mathbb{Z}}$  are  $\text{PGL}_2(\mathbb{Q}_p)$ -equivalent but not  $\text{PGL}_2(\mathbb{Z}_p)$ -equivalent, then by replacing  $f$  with a form that is  $\text{PGL}_2(\mathbb{Z}_p)$ -equivalent to it, we may assume that  $f = \gamma \cdot f'$ , where  $\gamma = \begin{pmatrix} p^n & \\ & p^{-n} \end{pmatrix}$ . It is then clear that  $f$  is not maximal, and therefore not strongly maximal, at  $p$ .

Next, we show that if  $f \in V_{\mathbb{Z}}$  is not  $\mathbb{Q}_p$ -soluble, then it is again not strongly maximal at  $p$ . First, if the discriminant of  $f \in V_{\mathbb{Z}_p}$  is prime to  $p$ , then  $f$  is  $\mathbb{Q}_p$ -soluble (see [21, Chapter 3.6]). Also, if the splitting type of  $f$  at  $p$  is  $(1^2 11)$  or  $(1^3 1)$ , then the reduction of  $f$  modulo  $p$  has a single root in  $\mathbb{P}^1(\mathbb{F}_p)$ , which lifts to a root in  $\mathbb{P}^1(\mathbb{Q}_p)$  by Hensel's Lemma. Thus  $f$  is  $\mathbb{Q}_p$ -soluble. If the splitting type of  $f$  over  $p$  is  $(2^2)$ ,  $(1^2 1^2)$ , or  $(1^4)$ , then  $f$  is overramified and is therefore not strongly maximal.

Thus it suffices to prove that if the splitting type of  $f$  at  $p$  is  $(1^2 2)$ , then  $f$  is  $\mathbb{Q}_p$ -soluble. If  $f \in V_{\mathbb{Z}_p}$  has splitting type  $(1^2 2)$ , then the reduction of  $f$  modulo  $p$  can be assumed to be of the form  $ax^2(x^2 - \bar{n}y^2)$ , where  $\bar{n}$  is a nonresidue modulo  $p$ . Hence we may assume that  $f$  is  $\text{PGL}_2(\mathbb{Q}_p)$ -equivalent to  $a(x^2 - kpy^2)(x^2 - ny^2)$ , where  $a, n, k \in \mathbb{Z}_p$ , the element  $n \in \mathbb{Z}_p$  is a nonresidue when reduced modulo  $p$ , and  $p \nmid a$ . If  $a$  is a square in  $\mathbb{Q}_p$ , then  $f(1, 0)$  is a square in  $\mathbb{Q}_p$  and we are done. So we may assume that  $a$  is a nonsquare. Now if  $p \nmid x$ , then  $x^2 - kp$  is a square in  $\mathbb{Q}_p$ ; so it suffices to prove the existence of  $x_0 \in \mathbb{F}_p^\times$  such that  $x_0^2 - \bar{n}$  is a nonresidue modulo  $p$ . Suppose this is not the case, and denote the residues in  $\mathbb{F}_p^\times$  by  $r_1, \dots, r_\ell$ , where  $\ell = (p-1)/2$ . If for all  $i$  we have that  $r_i - \bar{n}$  is a residue, then by the pigeonhole principle there exists  $c \leq \ell + 1$  and a residue  $r \in \mathbb{F}_p^\times$  such that  $r - c\bar{n} = r$ , which is a contradiction.  $\square$

### 5.4 Proofs of the main theorems (Theorems 1.1, 1.3, and 5.1)

We prove the following theorem, from which Theorem 5.1 will be seen to follow.

**Theorem 5.14** *Let  $F$  be a large set of elliptic curves. Then we have*

$$\lim_{X \rightarrow \infty} \frac{\sum_{\substack{E \in F \\ H'(E) < X}} (\#S_2(E) - 1)}{\sum_{\substack{E \in F^{\text{inv}} \\ H'(E) < X}} 1} = \zeta(2) \prod_p \frac{M_p(V, F)}{M_p(U_1, F)}. \quad (63)$$

**Proof:** We begin by counting the number of elliptic curves in  $F$  that have bounded height in terms of the volumes of the sets  $R_X^\pm$  which were defined in the proof of Proposition 2.10 by

$$R_X^+ = \{(i, j) \in \mathbb{R}^2 : |i| < X^{1/3}, |j| < 2X^{1/2}, 4i^3 - j^2 > 0\},$$

$$R_X^- = \{(i, j) \in \mathbb{R}^2 : |i| < X^{1/3}, |j| < 2X^{1/2}, 4i^3 - j^2 < 0\}.$$

**Lemma 5.15** *With the notation of Theorem 5.14, we have*

$$\sum_{\substack{E \in F \\ H'(E) < X}} 1 = [\text{Vol}(R_1^+) + \text{Vol}(R_1^-)] \cdot \prod_p M_p(U_1, F) \cdot X^{5/6} + o(X^{5/6}), \quad (64)$$

**Proof:** It was proved in Proposition 2.10 that the total number of points  $(I, J) \in \mathbb{Z} \times \mathbb{Z}$  such that  $H(I, J) < X$  is equal to  $\text{Vol}(R_X^+) + \text{Vol}(R_X^-)$ , up to an error of  $X^{1/2+\epsilon}$ . From (62) we see that the quantity  $M_p(U_1, F)$  is equal to the  $p$ -adic density of  $F_p^{\text{inv}}$  in  $\mathbb{Z}_p \times \mathbb{Z}_p$ . The lemma now follows from Proposition 3.17 along with the condition that  $F$  is a large set of elliptic curves. The proof is identical to that of Theorem 1.9 in §3.7.  $\square$

Now note that the numerator of the left hand side of the equation in Theorem 5.14 is equal to  $N(S^F; 2^{12}X)$ , the number of  $\text{PGL}_2(\mathbb{Q})$ -equivalence classes of locally soluble irreducible integral binary quartic forms having invariants  $2^4I$  and  $2^6J$ , such that  $(I, J) \in F^{\text{inv}}$  and  $H(2^4I, 2^6J) < 2^{12}X$ . By Propositions 5.12 and 5.13 we have, up to an error of  $o(X^{5/6})$ , that

$$\begin{aligned} N(S^F; 2^{12}X) &= N(V_{\mathbb{Z}}^{(0)} \cup V_{\mathbb{Z}}^{(2+)} \cup V_{\mathbb{Z}}^{(1)}; 2^{12}X) \cdot \prod_p \mu_p(S^F) \\ &= N(V_{\mathbb{Z}}^{(0)} \cup V_{\mathbb{Z}}^{(2+)} \cup V_{\mathbb{Z}}^{(1)}; 2^{12}X) \cdot \prod_p |2^{10}/3^3|_p \cdot M_p(V, F) \\ &= \frac{2}{27} \zeta(2) 2^{10} \left[ \left( \frac{1}{4} + \frac{1}{4} \right) \cdot \text{Vol}(R_1^+) + \frac{1}{2} \cdot \text{Vol}(R_1^-) \right] \cdot \prod_p |2^{10}/3^3|_p \cdot M_p(V, F) \cdot X^{5/6} \\ &= \zeta(2) [\text{Vol}(R_1^+) + \text{Vol}(R_1^-)] \cdot \prod_p M_p(V, F) \cdot X^{5/6}; \end{aligned} \quad (65)$$

the proof is again identical to that of Theorem 1.9. Taking the ratio of (65) and (64) now yields Theorem 5.14.  $\square$

To evaluate the right hand side of (63), we require the following fact (see [15, Lemma 3.1]):

**Lemma 5.16** *Let  $E$  be an elliptic curve over  $\mathbb{Q}_p$ . Then*

$$\#(E(\mathbb{Q}_p)/2E(\mathbb{Q}_p)) = \begin{cases} \#E[2](\mathbb{Q}_p) & \text{if } p \neq 2; \\ 2 \cdot \#E[2](\mathbb{Q}_p) & \text{if } p = 2; \end{cases}$$

By Proposition 5.12 and (62), the right hand side of (63) is equal to

$$\zeta(2) \prod_p \left( 1 - \frac{1}{p^2} \right) \frac{\prod_p \int_{(I,J) \in F_p^{\text{inv}}} \sum_{\sigma \in E_{I,J}/2E_{I,J}} \frac{1}{\#E[2](\mathbb{Q}_p)} dIdJ}{\prod_p \int_{(I,J) \in F_p^{\text{inv}}} dIdJ}$$

which is then equal to 2 by Lemma 5.16. We have proven Theorem 5.1 (and thus also Theorems 1.1 and 1.3).



## References

- [1] A. M. Baily, On the density of discriminants of quartic fields, *J. reine angew. Math.* **315** (1980), 190–210.
- [2] K. Belabas, M. Bhargava, and C. Pomerance, Error terms for the Davenport-Heilbronn theorems, *Duke Math. J.* **153** (2010), 173–210.
- [3] M. Bhargava, Higher composition laws I: A new view on Gauss composition, and quadratic generalizations, *Ann. of Math.* **159** (2004), no. 1, 217–250.
- [4] M. Bhargava, Higher composition laws III: The parametrization of quartic rings, *Ann. of Math.* **159** (2004) 1329–1360.
- [5] M. Bhargava, The density of discriminants of quartic rings and fields, *Ann. of Math.* **162**, 1031–1063.
- [6] M. Bhargava, The density of discriminants of quintic rings and fields, *Ann. of Math.*, to appear.
- [7] M. Bhargava, Mass formulae for extensions of local fields, and conjectures on the density of number field discriminants, *Int. Math. Res. Not.*, IMRN 2007, no. 17, Art. ID rnm052, 20 pp.
- [8] M. Bhargava, J. Hanke, and A. Shankar, The average number of 2-torsion class group elements in  $n$ -monogenic fields, in progress.
- [9] M. Bhargava, A. Shankar, and J. Tsimerman, On the Davenport-Heilbronn theorems, and second order terms, submitted.
- [10] M. Bhargava and A. Shnidman, On the number of cubic orders of bounded discriminant having automorphic group  $C_3$ , and related problems, in preparation.
- [11] B. J. Birch and J. R. Merriman, Finiteness theorems for binary forms with given discriminant, *Proc. London Math. Soc. (3)* **24** (1972), 385–394.
- [12] B. J. Birch and H. P. F. Swinnerton-Dyer, Notes on elliptic curves. I. *J. Reine Angew. Math.* **212** 1963 7–25.
- [13] A. Borel and Harish-Chandra, Arithmetic subgroups of algebraic groups, *Ann. of Math.* **75** (1962), 485–535.
- [14] A. Brumer, The average rank of elliptic curves. I. *Invent. Math.* **109** (1992), no. 3, 445–472.
- [15] A. Brumer and K. Kramer, The rank of elliptic curves, *Duke Math J.* **44** (1977), no. 4, 715–743.
- [16] H. Cohen and H. W. Lenstra, Heuristics on class groups of number fields, Number theory (Noordwijkerhout, 1983), 33–62, *Lecture Notes in Math.* **1068**, Springer, Berlin, 1984.
- [17] H. Cohen and J. Martinet, Étude heuristique des groupes de classes des corps de nombres, *J. Reine Angew. Math.* **404** (1990), 39–76.
- [18] H. Cohen and J. Martinet, Heuristics on class groups: some good primes are not too good, *Math. Comp.* **63** (1994), no. 207, 329–334.
- [19] J. E. Cremona, M. Stoll, Minimal models for 2-coverings of elliptic curves. *LMS J. Comput. Math.* **5** (2002), 220–243 (electronic).
- [20] J. E. Cremona, Reduction of binary cubic and quartic forms, *LMS J. Comput. Math.* , 64–94.
- [21] J. E. Cremona, *Algorithms for modular elliptic curves*, 2nd edn., Cambridge University Press, 1997.
- [22] H. Davenport, On a principle of Lipshitz, *J. London Math. Soc.* **26** (1951), 179–183. Corrigendum: "On a principle of Lipschitz", *J. London Math. Soc.* **39** (1964), 580.

- [23] H. Davenport, On the class-number of binary cubic forms I and II, *J. London Math. Soc.* **26** (1951), 183–198.
- [24] H. Davenport and H. Heilbronn, On the density of discriminants of cubic fields II, *Proc. Roy. Soc. London Ser. A* **322** (1971), no. 1551, 405–420.
- [25] B. N. Delone, Über die Darstellung der Zahlen durch die binäre kubischen Formen von negativer Diskriminante, *Math. Z.* **31** (1930), 1–26.
- [26] B. N. Delone and D. K. Faddeev, *The theory of irrationalities of the third degree*, AMS Translations of Mathematical Monographs **10**, 1964.
- [27] J. H. Evertse, On the representation of integers by binary cubic forms of positive discriminant, *Invent. Math.* **73** (1983), no. 1, 117–138.
- [28] C. F. Gauss, *Disquisitiones Arithmeticae*, 1801.
- [29] D. Goldfeld, Conjectures on elliptic curves over quadratic fields, *Number theory, Carbondale 1979 (Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, Ill., 1979)*, pp. 108–118, Lecture Notes in Math., **751**, Springer, Berlin, 1979.
- [30] D. R. Heath-Brown, The average analytic rank of elliptic curves. *Duke Math. J.* **122** (2004), no. 3, 591–623.
- [31] H. Heilbronn, On the 2-classgroup of cubic fields, *Studies in Pure Math.* (1971), 117–119.
- [32] N. M. Katz, and P. Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy*, American Mathematical Society Colloquium Publications **45**, American Mathematical Society, Providence, RI, 1999.
- [33] K. S. Kedlaya, Mass formulas for local Galois representations. With an appendix by Daniel Gulotta. *Int. Math. Res. Not. IMRN* 2007, no. 17, Art. ID rnm021, 26 pp.
- [34] A. W. Knap, Lie groups beyond an introduction. Second edition. Progress in Mathematics **140**, Birkhäuser Boston, Inc., Boston, MA, 2002. xviii+812 pp.
- [35] G. Malle, The totally real primitive number fields of discriminant at most  $10^9$ , ANTS VII, Berlin, 114–123, *Lecture Notes in Comput. Sci.* **4076**, Springer, Berlin, 2006.
- [36] F. Mertens, Ueber einige asymptotische Gesetze der Zahlentheorie, *J. reine angew. Math.* **77** (1874), 289–338.
- [37] J. Nakagawa, Orders of a Quartic Field, *Mem. Amer. Math. Soc.* **122** (1996), no. 583, A.M.S., Providence, RI.
- [38] J-P. Serre, Une “formule de masse” pour les extensions totalement ramifiées de degré donné d’un corps local, *C. R. Acad. Sci. Paris Sér. A-B* **286** (1978), no. 22, A1031–A1036.
- [39] C. L. Siegel, The average measure of quadratic forms with given determinant and signature, *Ann. of Math. (2)* **45** (1944), 667–685.
- [40] J. H. Silverman, The arithmetic of elliptic curves, Second edition, *Graduate Texts in Mathematics* **106**, Springer, Dordrecht, 2009.
- [41] M. Wood, Moduli spaces for rings and ideals, Ph.D. Thesis, Princeton University, June 2009.
- [42] M. P. Young, Low-lying zeros of families of elliptic curves. *J. Amer. Math. Soc.* **19** (2006), no. 1, 205–250.